CHAPTER 14

ENSURING INTEGRITY AND AVAILABILITY

After reading this chapter and completing the exercises, you will be able to:

- Identify the characteristics of a network that keep data safe from loss or damage
- Protect an enterprise-wide network from viruses
- Explain network- and system-level fault-tolerance techniques
- Discuss issues related to network backup and recovery strategies
- Describe the components of a useful disaster recovery plan

ON THE JOB

I work at a weekly local newspaper with a circulation of about 50,000. Although I'm not an IT professional, I usually end up taking care of our computers, answering technical questions, and picking consultants to help with our network. Our internal network is small, with about 30 workstations connected over Ethernet. But our connection to the outside world—the Web, e-mail, printers, and other news agencies—is really our lifeblood. Without our WAN connections, we could not produce a paper.

A few years ago I hired a consultant to make sure our WAN connections were optimized. He decided we needed a DSL link to a regional DSL provider. The DSL provider also supplied Web hosting and e-mail services for us, all for an attractive price. This worked well for a long time, which means that I didn't even have to think about the WAN. But one day, without notice, our DSL provider went out of business. Suddenly we lost all contact with the outside world. We could not retrieve stories from our freelance writers, nor could we issue files to our printer. In fact, the staff couldn't even communicate electronically with each other. And we had a paper to get out in two days.

Needless to say, I did not call the same consultant who arranged for our original WAN installation. Instead, I called a larger network consulting firm in town that had experience with high availability and fault-tolerant networking. They quickly provided our newspaper with an emergency WAN link, in order to meet our immediate deadlines. Then they taught us how to keep data and connections always available. Among other things, we now have two connections to the Internet, each of which uses a different ISP.

Paige DeYoung
Cormier Consolidated News

As networks take on more of the burden of transporting and storing a day's work, you need to pay increasing attention to the risks involved. You can never assume that data are safe on the network until you have taken explicit measures to protect the information. In this book, you have learned about the architecture of a robust enterprise-wide network as well as hardware, network operating systems, and network troubleshooting. But all the best equipment and software cannot ensure that server hard drives will never fail or that a malicious employee won't sabotage your network.

The topic of protecting data covers a lot of ground, from fault-tolerant servers to security cameras in the computer room. This chapter provides a broad overview of measures that you can take to ensure that your data remain safe. Undoubtedly, these issues will continue to evolve quickly as networks become more open and ubiquitous. If you are interested in specializing in fault tolerance, for example, you can read entire books on the topic. The far-reaching topic of network security is covered in the next chapter.

WHAT ARE INTEGRITY AND AVAILABILITY?

Before learning how to ensure integrity and availability, you should fully understand what these terms mean. **Integrity** refers to the soundness of a network's programs, data, services, devices, and connections. To ensure a network's integrity, you must protect it from anything that might render it unusable. Closely related to the concept of integrity is availability. **Availability** of a file or system refers to how consistently and reliably it can be accessed by authorized personnel. For example, a server that allows staff to log on and use its programs and data 99.99% of the time is considered to be highly available. To ensure availability, you need not only a well-planned and well-configured network, but also data backups, redundant devices, and protection from malicious intruders who could potentially immobilize the network.

A number of phenomena may compromise both integrity and availability, including security breaches, natural disasters (such as tornadoes, floods, hurricanes, and ice storms), malicious intruders, power flaws, and human error. Every network administrator should consider these possibilities when designing a sound network. You can readily imagine the importance of integrity and availability of data in a hospital, for example, where the network not only stores patient records but also provides quick medical reference material, video displays for surgical cameras, and perhaps even control of critical care monitors.

Even if you don't have sophisticated hardware and software to address availability and integrity, as network administrator you can and should take several precautions. This section will remind you of common-sense approaches to data integrity and availability, such as properly restricting file access and developing an enterprise-wide security policy. Later in this chapter, you will learn about more specific or formal (and potentially more expensive) approaches to data protection.

If you have ever supported computer users, you know that they sometimes unintentionally harm their own data, applications, software configurations, or even hardware. Networks may also be intentionally harmed by users unless network administrators take precautionary measures and pay regular, close attention to systems and networks so as to protect them. Although you can't predict every type of vulnerability, you can take measures to guard against most damaging events. Following are some general guidelines for protecting your network:

- *Prevent anyone other than a network administrator from opening or changing the system files.* Pay attention to the rights assigned to regular users (including the groups "users" or "everyone"). The use of rights to restrict network access to servers will be discussed in depth in Chapter 15. For now, bear in mind that the worst consequence of applying overly stringent file restrictions is a temporary inconvenience to a few users. In contrast, the worst consequence of applying overly lenient file restrictions could be a network disaster.
- *Monitor the network for unauthorized access or changes.* You can install programs that routinely check whether and when the files you've specified (for example, autoexec.ncf on a NetWare server) have changed. Such monitoring programs are typically inexpensive and easy to customize. They may even enable the system to page or e-mail you when a system file changes. In addition, you can monitor the network for unauthorized access to devices such as routers or switches. This practice, called **intrusion detection**, is described in more detail later in this chapter.
- *Record authorized system changes in a change management system.* In Chapters 12 and 13, you learned about the importance of change management. Recording system changes in a change management system will enable you and your colleagues to understand what's happening to your network and protect it from harm. For example, suppose that a Windows 2000 server hangs up when you attempt to restart it. Before launching into troubleshooting techniques that may create more problems and reduce the availability of the system, you could review the change management log. It might indicate that a colleague recently installed a new service pack. With this information in hand, you could focus on the service pack as the probable source of the problem.
- *Install redundant components.* The term **redundancy** refers to a situation in which more than one component is installed and ready to use for storing, processing, or transporting data. To maintain high availability, you should ensure that critical network elements, such as your WAN connection to the Internet or your single file server's hard disk, are redundant. Some types of redundancy require large investments, so your organization should weigh the risks of losing connectivity or data against the cost of adding expensive duplicate components such as data links or high-end servers.
- *Perform regular health checks on the network.* Prevention is the best weapon against network down time. By implementing a network monitoring program such as those discussed in Chapter 13, you can anticipate problems before they affect availability or integrity. For example, if your network monitor alerts you to rapidly rising utilization on a critical network segment, you can analyze the network to discover where the problem lies and perhaps fix it before it takes down the segment.

- *Monitor system performance, error logs, and the system log book regularly.* By keeping track of system errors and trends in performance, you have a better chance of correcting problems before they cause a hard disk failure and potentially damage your system files. By default, all network operating systems keep error logs. It's important that you know where these error logs reside on your server and understand how to interpret them.
- *Keep backups, boot disks, and emergency repair disks current and available.* If your file system or critical boot files become corrupted by a system crash, you can use the emergency or boot disks to recover the system. Otherwise, you may need to reinstall the software before you will be able to start the system. If you ever face the prospect of recovering from a system loss or disaster, you will need to recover in the quickest manner possible. For this effort, you will need not only backup devices, but also a backup strategy tailored to your environment.
- *Implement and enforce security and disaster recovery policies.* Everyone in your organization should know what he or she is allowed to do on the network. For example, if you decide that it's too risky for employees to download games off the Internet because of the potential for virus infection, you may inform them of a ban on downloading games. You might enforce this policy by restricting users' ability to create or change files (such as executable files) that are copied to the workstation during the downloading of games. Making such decisions and communicating them to staff should be part of your security policy. Likewise, everyone in your organization should be familiar with your disaster recovery plan, which should detail your strategy for bringing the network back to functionality in case of an unexpected failure. Although such policies take time to develop and may be difficult to enforce, they can directly affect your network's availability and integrity.

These measures are merely first steps to ensuring network integrity and availability, but they are essential. The following sections describe what types of policies, hardware, and software you can implement to achieve availability and integrity, beginning with virus detection and prevention.

VIRUSES

Strictly speaking, a **virus** is a program that replicates itself so as to infect more computers, either through network connections or through floppy disks passed among users. A virus may damage files or systems, or it may simply annoy users by flashing messages or pictures on the screen or by causing the computer to beep. In fact, some viruses cause no harm and can remain unnoticed on a system forever.

Many other unwanted and potentially destructive programs are mistakenly called viruses. For example, a program that disguises itself as something useful but actually harms your system is called a **Trojan horse**, after the famous wooden horse in which soldiers were hidden.

Because Trojan horses do not replicate themselves, they are not technically viruses. An example of a Trojan horse is an executable file that someone sends you over the Internet, promising that the executable will install a great new game, when in fact it reformats your hard disk.

In this section, you will learn about the different types of viruses and other malicious programs that may infect your network, their methods of distribution, and, most importantly, protection against them. Viruses can infect computers running any type of operating system—Macintosh, NetWare, Windows, or UNIX—at any time. As a network administrator, you must take measures to guard against them.

Types of Viruses

Many thousands of viruses exist, although only a relatively small number cause the majority of virus-related damage. Viruses can be classified into different categories based on where they reside on a computer and how they propagate themselves. Often, creators of viruses apply slight variations to their original viruses to make them undetectable by antivirus programs. The result is a host of related, albeit different viruses. The makers of antivirus software must then update their checking programs to recognize the new variations, and the virus creators may again alter their viruses to render them undetectable. This cycle continues, ad infinitum. No matter what their variation, all viruses belong to one of the categories described below:

- *Boot sector viruses*—The most common types of viruses, **boot sector viruses** reside on the boot sector of a floppy disk and become transferred to the partition sector or the DOS boot sector on a hard disk. The only way to infect a computer with a boot sector virus is to attempt to start the computer from an infected floppy disk. This event may happen unintentionally if a floppy disk is left in the drive when a machine starts.

  For example, one afternoon a colleague may give you a floppy disk with a spreadsheet that you need to edit and return to him. You put the floppy into your disk drive and open the spreadsheet file. So far, the virus in the floppy disk's boot sector has gone unnoticed. You begin to edit the spreadsheet, but get sidetracked by a critical file server problem. It's six o'clock by the time you have fixed the file server, and you're late for your evening cooking class, so you close all programs, turn off your machine, and rush out the door. The next morning, you switch on your machine and walk away to refill your coffee cup. Because you left the floppy disk in your disk drive, your computer attempts to start from the floppy disk drive. It loads the first sector into memory and executes it (normally, this sector contains a program written by Microsoft to load DOS or, if it can't find DOS on the disk, to tell you so). Because the floppy drive is infected with a boot sector virus, however, it executes the virus program instead. The virus installs itself on your computer's hard disk, replacing the hard disk's boot sector record. Until you disinfect your computer, the virus will propagate to every floppy disk to which you write information.

Boot sector viruses are very common in part because most users don't understand how they work, and because floppy disks are frequently passed from user to user without any virus checking. Examples of boot sector viruses include "Stoned," "Boot-437," "Goldbug," "Lilith," "Jerusalem," and "Cascade." The Stoned virus, for example, originated in New Zealand in 1988; since then, a multitude of variations on it have been distributed under different names. Its main symptom of infection is a message that appears upon starting the computer, announcing that "This PC is now stoned." In addition, boot sector viruses often make it impossible for the file system to access at least some of the workstation's files.

- *Macro viruses*—**Macro viruses** are newer types of viruses that take the form of a word-processing or spreadsheet program macro, which may be executed as the user works with a word-processing or spreadsheet program. Macro viruses were the first type of virus to infect data files rather than executable files. Because data files are more apt to be shared among users, and because macro viruses are typically easier to write than executable viruses, macro viruses have quickly become prevalent. Although the earliest versions of macro viruses proved annoying but not harmful, currently circulating macro viruses may threaten data files.

  Because macro viruses work under different applications, they can travel between computers that use different operating systems. For example, you might send a Microsoft Word document as an attachment to an e-mail message, or give it to someone on a floppy disk. If that document contains a macro virus, when the recipient opens the document, the macro runs, and all future documents created or saved by that program will be infected. Examples of macro viruses include "W97M/Ethan.A," "Laroux," "Trasher," "Caligula," and "Jedi." Symptoms of macro virus infection vary widely but may include missing options from application menus; damaged, changed, or missing data files; or strange pop-up messages that appear when you use an application such as Microsoft's Word or Excel.

- *File-infected viruses*—**File-infected viruses** attach themselves to executable files. When the infected executable file runs, the virus copies itself to memory. Later, the virus will attach itself to other executable files. Some file-infected viruses can attach themselves to other programs even while their "host" executable runs a process in the background, such as a printer service or screen saver program. Because they stay in memory while you continue to work on your computer, these viruses can have devastating consequences, infecting numerous programs and requiring you to not only disinfect your computer, but also reinstall virtually all software. Examples of file-infected viruses include "Tequila," "Concept," "Anxiety," "Tentacle," and "Cabanas." Symptoms of a virus infection may include damaged program files, inexplicable file size increases, changed icons for programs, strange messages that appear when you attempt to run a program, or the inability to run a program.

- *Network viruses*—**Network viruses** propagate themselves via network protocols, commands, messaging programs, and data links. Although all viruses could theoretically travel across network connections, network viruses are specially designed to take advantage of network vulnerabilities. For example, a network virus may attach itself to FTP transactions to and from your Web server. Another type of network virus may spread through Microsoft Exchange messages only.

  Because network access has become more sophisticated over the last decade, few network viruses have had the opportunity to thrive. Examples of network viruses include "Homer," "WDEF," and "Remote Explorer." Because network viruses are characterized by their transmission method, their symptoms may include almost any type of anomaly, ranging from strange pop-up messages to file damage.

- *Worms*—**Worms** are not technically viruses, but rather programs that run independently and travel between computers and across networks. They may be transmitted by any type of file transfer, including e-mail. Worms do not alter other programs in the same way that viruses do, but they may carry viruses. Because they can transport (and hide) viruses, you should be concerned about picking up worms when you exchange files from the Internet or through floppy disks. Examples of worms include "W32/Roach@MM," "SunOS/BoxPoison," and "W32/Mona." Symptoms of worm infection may include almost any type of anomaly, ranging from strange pop-up messages to file damage.

- *Trojan horse*—As mentioned earlier, a Trojan horse (sometimes simply called a "Trojan") is not actually a virus, but rather a program that claims to do something useful but instead harms the computer or system. Trojan horses range from being nuisances to causing significant system destruction. Most virus-checking programs will recognize known Trojan horses and eradicate them. The best way to guard against Trojan horses, however, is to refrain from downloading an executable file whose origins you can't confirm.

  Suppose, for example, that you needed to download a new driver for a NIC on your network. Rather than going to a generic "network support site" on the Internet, you should download the file from the NIC manufacturer's Web site. Most importantly, never run an executable file that has been sent to you over the Internet as an attachment to a mail message whose sender or origins you cannot verify.

  Examples of Trojan horses include "BackDoor-G2.svr," "*VBS/FreeLink@MM*," "Sadcase," "Perl-WSFT-Exploit," and "DOS/Blitz." One Trojan horse program, "Antigen," disguises itself as an antivirus program; when executed, it scans the computer's hard disk for personal information such as network IDs, passwords, and telephone numbers. It then compiles this information and mails it to a specific e-mail address.

Virus Characteristics

Viruses that belong to any of the preceding categories may have additional characteristics that make them harder to detect and eliminate. Some of these characteristics are discussed below:

- **Encryption**—Some viruses are encrypted to prevent detection. As you will learn in the following section, most virus-scanning software searches files for a recognizable string of characters that identify the virus. If the virus is encrypted, it may thwart the antivirus program's attempts to detect it.
- **Stealth**—Some viruses hide themselves to prevent detection. Typically, stealth viruses disguise themselves as legitimate programs or replace part of a legitimate program's code with their destructive code.
- **Polymorphism**—Polymorphic viruses change their characteristics (such as the arrangement of their bytes, size, and internal instructions) every time they are transferred to a new system, making them harder to identify. Some polymorphic viruses use complicated algorithms and incorporate nonsensical commands to achieve their changes. Polymorphic viruses are considered to be the most sophisticated and potentially dangerous type of virus.
- **Time-dependence**—Time-dependent viruses are programmed to activate on a particular date. These types of viruses, also known as "time bombs," can remain dormant and harmless until their activation date arrives. Like any other type of virus, time-dependent viruses may have destructive effects or may cause some innocuous event periodically. For example, viruses in the "Time" family cause a PC's speaker to beep approximately once per hour.

Hundreds of new viruses are unleashed on the world's computers each month. Although it is impossible to keep abreast of every virus in circulation, you should at least know where you can find out more information about viruses. An excellent resource for learning about new viruses, their characteristics, and ways to get rid of them is McAfee's Virus Information Library at *vil.mcafee.com/default.asp*.

Virus Protection

Now that you know about the different types of viruses, you may think that you can simply install a virus-scanning program on your network and move on to the next issue. In fact, virus protection involves more than just installing antivirus software. It requires choosing the most appropriate antivirus program for your environment, monitoring the network, continually updating the antivirus program, and educating users. In addition, you should draft and enforce an antivirus policy for your organization.

Antivirus Software

Even if a user doesn't immediately notice a virus on his or her system, the virus will generally leave evidence of itself, whether by changing the operation of the machine or by announcing its signature characteristics in the virus code. Although the latter can be detected only via antivirus software, users can typically detect the former changes without any special software. For example, you may suspect a virus on your system if any of the following symptoms appear:

- Unexplained increases in file sizes
- Programs (such as Microsoft Word) launching, running, or exiting more slowly than usual
- Unusual error messages appearing without probable cause
- Significant, unexpected loss of system memory
- Fluctuations in display quality

Often, however, you will not notice a virus until it has already damaged your files.

Although virus programmers have become more sophisticated in disguising their viruses (for example, using encryption and polymorphism), antivirus software programmers have kept pace with them. The antivirus software you choose for your network should at least perform the following functions:

- It should detect viruses through **signature scanning**, a comparison of a file's content with known virus signatures (that is, the unique identifying characteristics in the code) in a signature database. This signature database must be frequently updated so that the software can detect new viruses as they emerge. Updates can usually be downloaded from the antivirus software vendor's Web site.
- It should detect viruses through **integrity checking**, a method of comparing current characteristics of files and disks against an archived version of these characteristics to discover any changes. The most common example of integrity checking involves the use of a checksum, though this tactic may not prove effective against viruses with stealth capabilities.
- It should detect viruses by monitoring unexpected file changes or virus-like behaviors.
- It should receive regular updates and modifications from a centralized network console. The vendor should provide free upgrades on a regular (at least monthly) basis, plus technical support.
- It should consistently report only valid viruses, rather than reporting "false alarms." Scanning techniques that attempt to identify viruses by discovering "virus-like" behavior, also known as **heuristic scanning**, are the most fallible and most likely to emit false alarms. As you might imagine, using an antivirus package that detects more viruses than are actually present can be not only annoying, but also a waste of time.

Caution

Occasionally, shrink-wrapped, off-the-shelf software will ship with viruses on its disks. Therefore, it is always a good idea to scan authorized software from known sources just as you would scan software from unknown sources.

Your implementation of antivirus software will depend on your computing environment's needs. For example, you may use a desktop security program on every computer on the network that prevents users from copying executable files to their hard disks or to network drives. In this case, it may be unnecessary to implement a program that continually scans each machine; in fact, this approach may be undesirable because the continual scanning may adversely impact performance. On the other hand, if you are the network administrator for a student computer lab where potentially thousands of different users will bring their own disks for use on the computers, you will want to scan the machines thoroughly at least once a day and perhaps more often.

When installing antivirus software on a network, one of your most important decisions is where to put it. If you install antivirus software only on every desktop, you have addressed the most likely point of entry, but ignored the most important files that might be infected—those on the server. If the antivirus software resides on the server and checks every file and transaction, you will protect important files but slow your network performance considerably. Likewise, if you put antivirus software on firewalls and routers, your network will experience performance problems, bringing all network communication to a crawl. How can you find a balance between sufficient protection and minimal impact on performance? Depending on your network infrastructure, you may want to implement antivirus software that scans each desktop once daily, as well as scans new files on the e-mail server, as those locations are the most likely places for viruses to enter. You should also ensure that file servers are scanned regularly, although continual may be unnecessary.

Obviously, the antivirus package you choose should be compatible with your network and desktop operating systems. Popular antivirus packages include Network Associate's (McAfee's) VirusScan, Computer Associates' Innoculan AntiVirus, Norman Virus Control, and Symantec's (Norton's) AntiVirus.

Tip

In addition to using specialized antivirus software to guard against virus infection, you may find that your applications can help identify viruses. Microsoft's Word and Excel programs, for example, will warn you when you attempt to open a file that contains macros. You then have the option of disabling the macros (thereby preventing any macro viruses from working when you open the file) or allowing the macros to remain usable. In general, it's a good idea to disable the macros in a file that you have received from someone else, at least until after you have checked the file for viruses with your virus scanning software.

Antivirus Policies

Antivirus software alone will not keep your network safe from viruses. You also need to implement policies that limit the potential for users to introduce viruses to their workstations and to the network. The importance of these policies will increase as a network grows larger and more accessible and therefore becomes more susceptible to viruses.

To understand why, think of a day-care center attended by only two children with one adult supervising. These three people will bring and share whatever germs they have encountered outside the day-care center; any one person could catch the germs of the other two. If the day-care center houses 20 children and seven adults, however, the number of germs that people may pass to each other multiplies. Now any single person could catch the germs of 26 others. Similarly, a network with 1,000 users, each of whom might bring floppy disks from home and download files off the Web, inherently carries a greater risk of virus infection than a network serving only 10 users.

Because most computer viruses can be prevented by the application of a little technology and a little intelligence, it's important that all network users understand how to prevent viruses. An antivirus policy should provide rules for using antivirus software and policies for installing programs, sharing files, and using floppy disks. Furthermore, it should be authorized and supported by the organization's management, and sanctions should by outlined for disobeying the policy. Some good, general guidelines for an antivirus policy are as follows:

- Every computer in an organization should be equipped with virus detection and cleaning software that regularly scans for viruses. This software should be centrally distributed and updated to stay current with newly released viruses.
- Users should not be allowed to alter or disable the antivirus software.
- Users should know what to do in case their antivirus program detects a virus. For example, you might recommend that the user not continue working on his or her computer, but instead call the help desk and receive assistance in disinfecting the system.
- Every organization should have an antivirus team that focuses on maintaining the antivirus measures in place. This team would be responsible for choosing antivirus software, keeping the software updated, educating users, and responding in case of a significant virus outbreak.
- Users should be prohibited from installing any unauthorized software on their systems. This edict may seem extreme, but in fact users bringing programs (especially games) on disk from home are the most common source of viruses. If your organization permits game playing, you might institute a policy in which every game must be first checked for viruses and then installed on a user's system by a technician.
- Organizations should impose penalties on users who do not follow the antivirus policy.

When drafting an antivirus policy, bear in mind that these measures are not meant to restrict users' freedom, but rather to protect the network from serious damage and expensive down time. Explain to users that the antivirus policy protects their own data as well as critical system files. If possible, automate the antivirus software installation and operation so that users barely notice its presence. Do not rely on users to run their antivirus software each time they insert a disk or download a new program, because they will quickly forget to do so.

Virus Hoaxes

As in any other community, rumors sometimes spread through the Internet user community. One type of rumor consists of a false alert about a dangerous, new virus that could cause serious damage to your workstation. Such an alert is known as a **virus hoax**.Virus hoaxes usually have no realistic basis and should be ignored, as they merely attempt to create panic. Sometimes the origins of virus hoaxes can be traced (for example, the famous virus hoax, "GoodTimes," was traced to students at Swarthmore College), but often their sources remain anonymous.

A typical example of a virus hoax is one called "It Takes Guts to Say 'Jesus'," in which the body of the message says the following:

> VIRUS WARNING !!!!!!!
> If you receive an e-mail titled "It Takes Guts to Say 'Jesus'," DO NOT open it. It will erase everything on your hard drive. Forward this letter to as many people as you can. This is a new, very malicious virus and not many people know about it. This information was announced yesterday morning from IBM; please share it with people who might access the Internet.

Notice that the hoax warns that the virus will erase everything on your hard drive. In fact, no current virus can erase your hard drive when you merely open an infected e-mail message. Only an executable file, such as a Trojan horse, can accomplish this damage.Virus hoaxes also typically demand that you pass the alert to everyone in your Internet address book, thus propagating the rumor.

Virtually the only way to decide whether a message that warns about a virus is a hoax is to look it up on a Web page that lists virus hoaxes. A good resource for verifying virus hoaxes is *www.icsalabs.com/html/communities/antivirus/hoaxes.stml*. This Web site also allows you to learn more about the phenomenon of virus hoaxes.

If you or your colleagues receive a virus hoax, simply ignore it. Educate your colleagues to do the same, explaining why virus hoaxes should not cause alarm. Remember, however, that even a virus hoax message could potentially contain an *attached* file that does cause damage if executed. Once again, the best policy is to refrain from running any program whose origins you cannot verify.

Fault Tolerance

Besides guarding against viruses, another key factor in maintaining the availability and integrity of data is fault tolerance. **Fault tolerance** is the capacity for a system to continue performing despite an unexpected hardware or software malfunction. Before you can understand the issues related to fault tolerance, you must recognize the difference between failures and faults as they apply to networks. In broad terms, a **failure** is a deviation from a specified level of system performance for a given period of time. In other words, a failure occurs when something doesn't work as promised or as planned. For example, if your car breaks down on the highway, you can consider the breakdown to be a failure. A **fault**, on the other hand, involves the malfunction of one component of a system. A fault can result in a failure. For example, the fault that caused your car to break down might be a leaking water pump. The goal of fault-tolerant systems is to prevent faults from progressing to failures.

Fault tolerance can be achieved in varying degrees, with the optimal level of fault tolerance for a system depending on how critical its services and files are to productivity. At the highest level of fault tolerance, a system would remain unaffected by a drastic problem, such as a power failure. For example, an uninterruptible power supply (UPS) or a gas-powered generator that supplies electricity to a server despite a city-wide power failure provides high fault tolerance.

In addition to using alternative power sources, fault tolerance can be achieved through mirroring. When two servers mirror each other, they can quickly take over for their partner if it should fail. The process of one component immediately assuming the duties of an identical component is known as automatic **fail-over**. Even if one server's NIC fails, for example, fail-over ensures that the other server can automatically handle the first server's responsibilities. In highly fault-tolerant schemes, network users will not even recognize that a problem has occurred. In a moderately fault-tolerant system, on the other hand, users may have to endure brief service outages. An example of a moderately fault-tolerant system is one in which two servers mirror each other's data, but require a network administrator to intervene and switch users from one server to the other.

An excellent way to achieve fault tolerance is to provide duplicate, or redundant, elements to compensate for faults in critical components. You can implement redundancy for servers, cabling, routers, hubs, gateways, NICs, hard disks, power supplies, and other components. The most common type of network redundancy is data backup. **Hard disk redundancy**, called **RAID (Redundant Array of Inexpensive Disks)**, represents a sophisticated means for dynamically replicating data over several physical hard drives. These and other fault-tolerant techniques are discussed in more depth in later sections, which are ordered according to the layer of the OSI Model to which they correspond, from the Physical layer to the Application layer.

To assess the fault tolerance of your network, you must identify any single point of failure—that is, a point on the network where, if a fault occurs, the transfer of data may break down without possibility of an automatic recovery. For instance, if a LAN in your home consists of three PCs, each of which is connected to a hub and a file server in the basement, your

LAN has several single points of failure: the connection between the hub and the file server; the hub itself; each of the hub's ports; the electrical connection that powers the hub; the electrical connection that powers the file server; the file server's NIC, fan, hard disk, memory, and processor; and—depending on the criticality of each PC—potentially all of their connections and components.

Redundancy is intended to eliminate single points of failure. If your network cannot tolerate any down time, you must consider redundancy for power, cabling, hard disks, NICs, data links, and any other components that might halt operations if they suffer a fault. As you can imagine, complete redundancy is expensive. Therefore, you must understand not only where your network's single points of failure exist, but also how their malfunctioning might affect the network.

Environment

As you consider sophisticated fault-tolerance techniques for servers, routers, and WAN links, remember to analyze the physical environment in which your devices operate. Part of your data protection plan involves protecting your network from excessive heat or moisture, break-ins, and natural disasters. In the case of natural disasters, the best approach is to store data backups in a location other than where your servers reside.

In addition, you should make sure that your telecommunications closets and equipment rooms are air-conditioned and maintained at a constant humidity, according to the hardware manufacturer's recommendations. You can purchase temperature and humidity monitors that trip alarms if specified limits are exceeded. These monitors can prove very useful because the temperature can rise rapidly in a room full of equipment, causing overheated equipment to fail.

Power

No matter where you live, you have probably experienced a complete loss of power (a blackout) or a temporary dimming of lights (a brownout). Such fluctuations in power are frequently caused by forces of nature such as hurricanes, tornadoes, or ice storms. They may also occur when a utility company performs maintenance or construction tasks. The following section describes the types of power fluctuations for which network administrators should prepare. The next two sections describe alternative power sources, such as a UPS (uninterruptible power supply) or electrical generator, that can compensate for these flaws.

Power Flaws

Whatever the cause, networks cannot tolerate power loss or less than optimal power. The following list describes power flaws that can damage your equipment:

- **Surge**—A momentary increase in voltage due to distant lightning strikes or electrical problems. Surges may last only a few thousandths of a second, but several surges can degrade a computer's power supply. Surges are common. Indeed, without a surge protector, systems will be subjected to multiple surges each year.

- **Line noise**—A fluctuation in voltage levels caused by other devices on the network or electromagnetic interference. Some line noise is unavoidable, but excessive line noise may cause a power supply to malfunction, immediately corrupting program or data files and gradually damaging motherboards and other computer circuits. When you turn on fluorescent lights or a laser printer and the lights dim, you have probably introduced noise into the electrical system. If you continue working on your computer during a lightning storm, your computer will be subject to line noise. Some UPSs guard against line noise, and any critical system should have this type of protection.
- **Brownout**—A momentary decrease in voltage; also known as a **sag**. An overtaxed electrical system may cause brownouts, which you may recognize in your home as a dimming of the lights. Such decreases in voltage can cause significant problems for computer devices. Most UPSs guard against brownouts.
- **Blackout**—A complete power loss. A blackout may or may not cause significant damage to your network. If you are performing a network operating system upgrade when a blackout occurs and you have not protected the server, its network operating system may be damaged so completely that the server will not restart and its operating system must be reinstalled from scratch. If the file server is idle when a blackout occurs, however, it may recover very easily. All UPSs are designed to compensate for blackouts, but how quickly and completely and for how long will depend on the particular unit. To handle extended blackouts or to support a building full of computers, you will need something more powerful than a UPS, such as a gas- or diesel-powered electrical generator.

Each of these power problems can adversely affect network devices and their availability. Not surprisingly then, network administrators must spend a great deal of money and time ensuring that power remains available and problem-free. The following sections describe devices and ways of dealing with unstable power.

Uninterruptible Power Supply (UPS)

A popular way to ensure that a network device does not lose power is to install an **uninterruptible power supply (UPS)**. A UPS is a battery-operated power source directly attached to one or more devices and to a power supply (such as a wall outlet), which prevents undesired features of the wall outlet's A/C power from harming the device or interrupting its services.

UPSs vary widely in the type of power aberrations they can rectify, the length of time for which they can provide power, and the number of devices they can support. Of course, they also vary widely in price. Some UPSs are intended for home use, designed to merely keep your PC running long enough for you to properly shut it down in case of a blackout. Other UPSs perform sophisticated operations such as line conditioning, power supply monitoring, and error notification. The type of UPS you choose will depend on your budget, the number and size of your systems, and the critical nature of those systems.

UPSs are classified into two general categories: standby and online. A **standby UPS** provides continuous voltage to a device by switching virtually instantaneously to the battery when it detects a loss of power from the wall outlet. Upon restoration of the power, the standby UPS switches the device back to using A/C power again. One problem exists with standby UPSs: in the brief amount of time that it takes the UPS to discover that power from the wall outlet has faltered, a sensitive device (such as a server) may have already detected the power loss and shut down or restarted. Technically, a standby UPS doesn't provide continuous power; for this reason, it is sometimes called an "offline" UPS. Nevertheless, standby UPSs may prove adequate even for critical network devices such as servers, routers, and gateways. They cost significantly less than online UPSs. Figure 14-1 depicts a standby UPS.

Figure 14-1 Standby UPSs

An **online UPS** uses the A/C power from the wall outlet to continuously charge its battery, while providing power to a network device through its battery. In other words, a server connected to an online UPS always relies on the UPS battery for its electricity. An online UPS offers the best kind of power redundancy available. Because the server never needs to switch from the wall outlet's power to the UPS's power, there is no risk of momentarily losing service. Also, because the UPS always provides the power, it can deal with noise, surges, and sags before the power reaches the attached device. As you can imagine, online UPSs are much more expensive than standby UPSs. Figure 14-2 shows an online UPS.

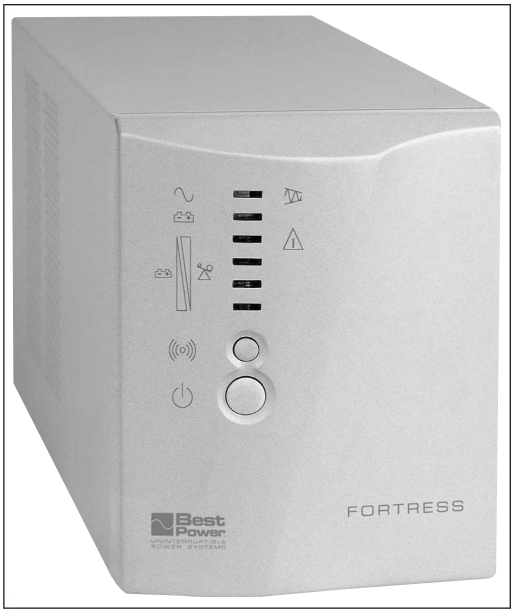


Figure 14-2 An online UPS

How do you decide which UPS is right for your network? You must consider a number of factors:

- *Amount of power needed*—The more power required by your device, the more powerful the UPS needed. Suppose that your organization decides to cut costs and purchase a UPS that cannot supply the amount of power required by a device. If the power to your building ever fails, this UPS will not support your device—you might as well have not installed any UPS.

  Electrical power is measured in volt-amps. A **volt-amp (VA)** is the product of the voltage and current (measured in amps) of the electricity on a line. To determine approximately how many VAs your device requires, you can use the following conversion: 1.4 volt-amps = 1 watt (W). A desktop computer, for example, may use a 200 W power supply and therefore require a UPS capable of at least 280 VA to keep the CPU running in case of a blackout. If you want backup power for your entire home office, however, you must account for the power needs for your monitor and any peripherals, such as printers, when purchasing a UPS. A medium-sized server with a monitor and external tape drive may use 402 W, thus requiring a UPS capable of providing at least 562 VA power. Determining your power needs can prove a challenge. Not only do you have to account for your existing equipment, but you should also consider how you might upgrade the supported device over the next several years. For example, you may purchase a server with only 4 GB of hard disk space, but plan to add 24 GB next year. When you upgrade the hard disk, you may also need to upgrade the UPS. Before you spend thousands of dollars on a UPS, consult with your equipment manufacturer to obtain its recommendations on power needs.

- *Period of time to keep a device running*—Most UPSs are rated to support a device for 15 to 20 minutes. The longer you anticipate needing a UPS to power your device, the more powerful your UPS must be. For example, the medium-sized server that could rely on a 574 VA UPS to remain functional for 20 minutes would need a 1100 VA server to remain functional for 90 minutes. To determine how long your device might require power from a UPS, consider the length of your typical power outages. If you live in an area that frequently suffers severe thunderstorms, you might want to purchase a higher-capacity UPS to cover longer outages.
- *Line conditioning*—Any UPS used on a network device should also offer surge suppression to protect against surges and line conditioning, or filtering, to guard against line noise. Line conditioners and UPS units include special noise filters that remove line noise. The manufacturer's technical specifications should indicate the amount of filtration required for each UPS. Noise suppression is expressed in decibel levels (dB) at a specific frequency (KHz or MHz). The higher the decibel level, the greater the protection.
- *Cost*—Prices for good UPSs vary widely, depending on the unit's size and extra features. A relatively small UPS that can power one server for 5 to 10 minutes might cost between $50 and $300. A large UPS that can power a sophisticated router for 10 to 20 minutes might cost between $200 and $3,000. On a critical system, however, you should not try to cut costs by buying an off-brand, potentially unreliable, or weak UPS.

As with other large purchases, you should research several UPS manufacturers and their products before reaching a decision. Also ensure that the manufacturer provides a warranty and lets you test the UPS with your equipment. It's important to try out the UPS with your equipment to ensure that it will satisfy your needs. Popular UPS manufacturers are APC, Best, Deltec, MGE, and Tripp Lite.

Generators

If your organization cannot withstand a power loss of any duration, either because of its computer services or other electrical needs, you might consider investing in an electrical generator for your building. Generators can be powered by diesel, liquid propane gas, natural gas, or steam. Although they do not provide surge protection, generators do provide clean (free from noise) electricity.

As when choosing a UPS, you should calculate your organization's crucial electrical demands to determine what size of generator you need. You should also estimate how long the generator may be required to power your building. Gas or diesel generators may cost between $10,000 and $3,000,000 (for the largest industrial types). Alternatively, you can rent electrical generators. To find out more about options for renting or purchasing generators in your area, contact your local electrical utility.

Topology

You have read about topology and architecture fault tolerance in previous chapters of this book. In Chapter 5, you learned about a variety of physical network topologies: star, ring, bus, mesh, and hybrid. Recall that each of these topologies inherently assumes certain advantages and disadvantages, and you need to assess your network's needs before designing your data links.

A mesh topology offers the best fault tolerance. To refresh your memory, a mesh network is one in which nodes are connected either directly or indirectly by multiple pathways. Figure 14-3 depicts a fully meshed network.

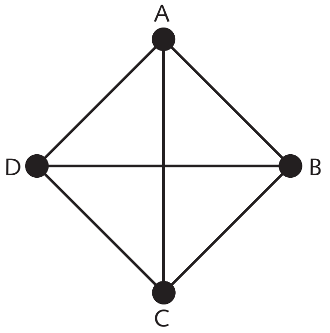


Figure 14-3 A fully meshed network

In a mesh topology, data can travel over multiple paths from any one point to another. For example, if the direct link between point A and point B in Figure 14-3 becomes severed, data can be rerouted automatically from point A to point C and then to point B. Alternatively, it may be rerouted from point A to point D to point B, and so on. You can see that a fully meshed network provides multiple redundancies and therefore greater fault tolerance than a network with a single redundancy.

Figure 14-4 illustrates a network that contains single redundancy. In this example, if one link between point A and point B becomes severed, data can automatically be rerouted over the second link. If the link between point A and point B and the link between point A and point C are both severed, however, the network will suffer a failure.

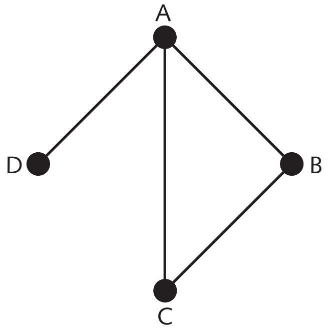


Figure 14-4 A network with one redundant connection

The physical media you use may also offer redundancy. Recall from Chapter 7 that a SONET ring can easily recover from a fault in one of its links because it forms a ring, as pictured in Figure 14-5. In this example, if the outer SONET link between point A and point B becomes severed, data can circumvent the fault to move between the two points.

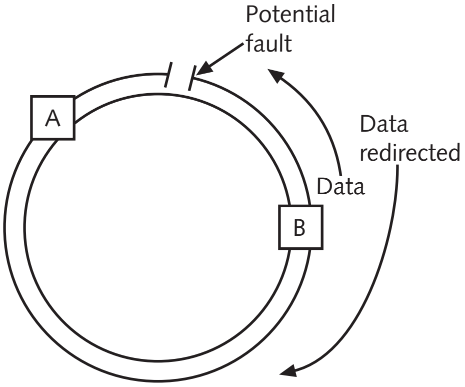


Figure 14-5 A self-healing SONET ring

Mesh topologies and SONET rings are good choices for highly available LANs and WANs. But what about connections to the Internet? Or data backup connections? You may need to establish more than one of these types of links.

As an example, imagine that you work for a data services firm called PayNTime that processes payroll checks for a large oil company in the Houston area. Every day you receive updated payroll information over a T1 link from your client, and every Thursday PayNTime compiles this information and then cuts 2,000 checks that you ship overnight to the client's headquarters. What would happen if the T1 link between PayNTime and the oil company suffered damage in a flood and became unusable on a Thursday morning? How would you ensure that the employees received their pay? If no redundant link to the oil company existed, you would probably need to gather and input the data into your system at least partially by hand. Even then, chances are that you wouldn't process the payroll checks in time to be shipped overnight.

In this type of situation, you would want a duplicate connection between PayNTime and the oil company's site. You might contract with two different service carriers to ensure the redundancy. Alternatively, you might arrange with one service carrier to provide two redundant routes. However you provide redundancy in your network topology, you should make sure that the critical data transactions can follow more than one possible path from source to target.

Redundancy in your network offers the advantage of reducing the risk of losing functionality, and potentially profits, from a network fault. As you might guess, however, the disadvantage of redundancy is its cost. If you subscribed to two different service providers for two T1 links in the PayNTime example, you would probably double your monthly leasing costs of approximately $1,000. Multiply that amount times 12 months, and then times the number of clients for which you need to provide redundancy—and the extra layers of protection quickly become expensive. Redundancy is like a homeowner's insurance policy: you may never need to use it, but if you don't get it, the cost can be much higher than your premiums. As a general rule, you should invest in connection redundancies where they are absolutely necessary.

Now suppose that PayNTime provides services not only to the oil company, but also to a temporary agency in the Houston area. Both links are critical because both companies need their payroll checks cut each week. With links to two customers, you may be able to take advantage of a T1 connection between the customers' sites to create a partially meshed network, as pictured in Figure 14-6. Now if the link between PayNTime and the oil firm suffers a fault, data can theoretically be rerouted through the temporary agency's connection.

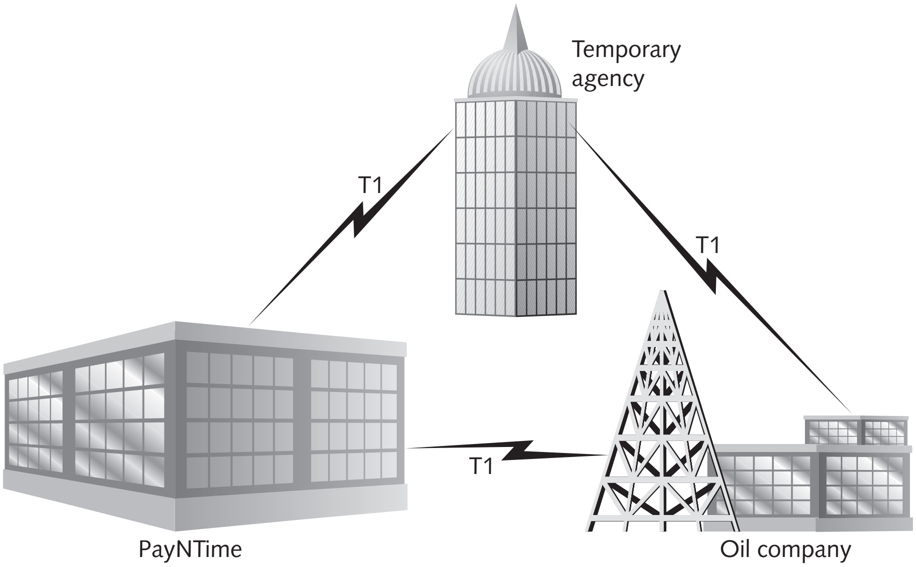


Figure 14-6 Redundancy between a firm and two customers

You may notice a problem with this scenario, however. What if the temporary agency doesn't want the oil company's transactions using its bandwidth, even in case of emergency? And what happens when the third and fourth customers are added to the network? To address concerns of capacity and scalability, you may want to consider partnering with an ISP and establishing secure VPNs with your clients. With a VPN, PayNTime could shift the costs of redundancy and network design to the service provider and concentrate on the task it does best—processing payroll. Figure 14-7 illustrates this type of arrangement.

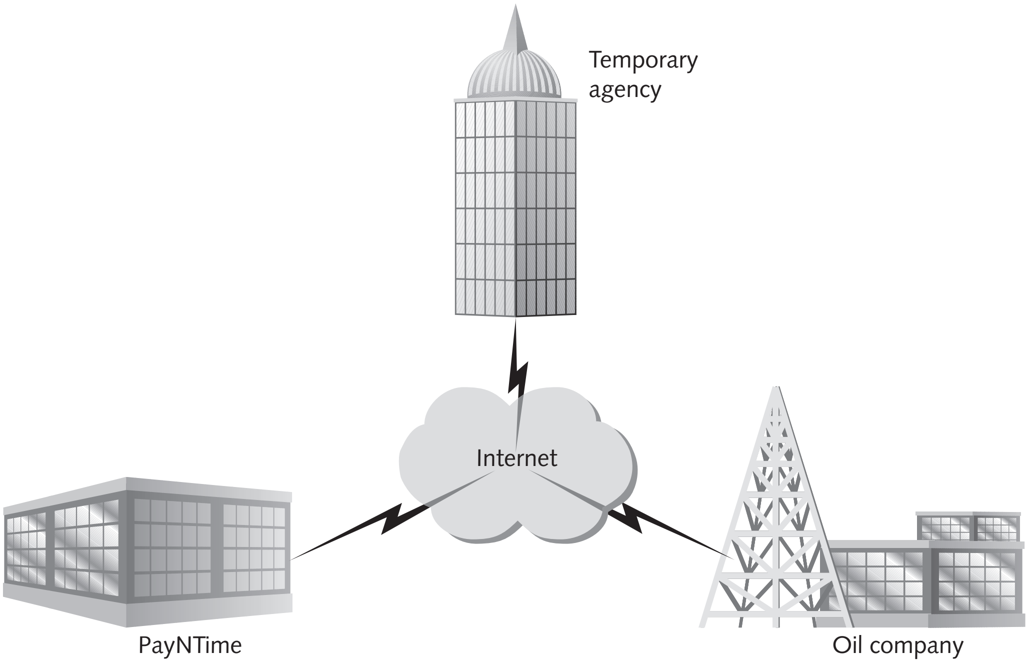


Figure 14-7 VPNs linking multiple customers

Connectivity

In the previous section, you learned the basics about providing fault tolerance in a LAN or WAN topology. But what about the devices that connect one segment of a LAN or WAN to another? What happens when they experience a fault? In Chapter 6, you learned how routers, bridges, hubs, and switches work. In Chapter 7, you saw how dedicated lines terminate at a customer's premises and in a service provider's data center. In this section, you will consider how to fundamentally increase the fault tolerance of connectivity devices and a LAN's or WAN's connecting links.

To understand how to increase the fault tolerance of not just the topology, but also the network's connectivity, let's return to the example of PayNTime. Suppose that the company's network administrator decides to establish a VPN agreement with a national ISP.

PayNTime's bandwidth analysis indicates that a T1 link will be sufficient to transport the data of five customers from the ISP's office to PayNTime's data room. Figure 14-8 provides a detailed representation of this arrangement.

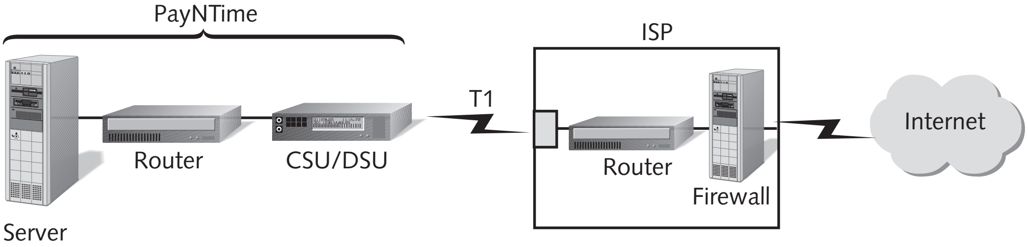


Figure 14-8 ISP connectivity

Notice the single points of failure in the arrangement depicted in Figure 14-8. As mentioned earlier, the T1 connection could incur a fault. In addition, any one of the routers, CSU/DSUs, or firewalls might suffer faults in their power supplies, NICs, or circuit boards. In a critical component such as a router or switch, high fault tolerance necessitates the use of redundant power supplies, cooling fans, interfaces, and I/O modules, all of which should ideally be hot swappable. The term **hot swappable** refers to identical components that automatically assume the functions of their counterpart if one suffers a fault. They are called hot swappable because they can be changed (or swapped) while a machine is still running (hot). In a sense, hot swappable components work like your kidneys. If one fails, the other will automatically assume all responsibility for filtering waste from the blood. In much the same way, if a router's processor fails, the redundant processor will automatically take over all data-processing functions. When you purchase switches or routers to support critical links, look for those that contain hot swappable components. As with other redundancy provisions, these features will add to the cost of your device purchase.

Purchasing connectivity devices does not address all faults that may occur on a WAN. In fact, faults may also affect the connecting links. For example, if you connect two offices with a dedicated T1 connection and the T1 fails, it doesn't matter whether your router has redundant NICs. The connection will still be down. Because a fault in the T1 link has the same effect as a bad T1 interface in a router, a fully redundant system might be a better option. Such a system is depicted in Figure 14-9.

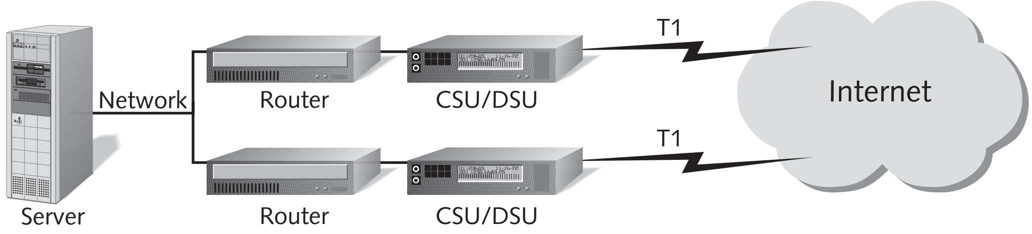


Figure 14-9 A fully redundant system

The preceding scenario utilizes the most expensive and reliable option for providing network redundancy for PayNTime. In addition, this solution allows for **load balancing**, or an automatic distribution of traffic over multiple links or processors to optimize response. Load balancing would maximize the throughput between PayNTime and its ISP because the aggregate traffic flowing between the two points could move over either T1 link, avoiding potential bottlenecks on a single T1 connection. Although one company might be willing to pay for such complete redundancy, another might prefer a less expensive solution. A less expensive redundancy option might be to use a dial-back WAN link. For example, a company that depends on a Frame Relay WAN might have an access server with an ISDN or 56 KB modem link that automatically dials the remote site when it detects a failure of the primary link.

Servers

As with other devices, you can make servers more fault-tolerant by supplying them with redundant components. Critical servers (such as those that perform user authentication for an entire LAN, or those that run important, enterprise-wide applications such as an electronic catalog in a library) often contain redundant NICs, processors, and hard disks. These redundant components provide assurance that if one item fails, the entire system won't fail; at the same time, they enable load balancing.

For example, a server with two 100-Mbps NICs, such as the one pictured in Figure 14-10, may be receiving and transmitting traffic at a rate of 46 Mbps during a busy time of the day. With additional software provided by either the NIC manufacturer or a third party, the redundant NICs can work in tandem to distribute the load, ensuring that approximately half the data travels through the first NIC and half through the second. This approach improves response time for users accessing the server. If one NIC fails, the other NIC will automatically assume full responsibility for receiving and transmitting all data to and from the server. Although load balancing does not technically fall under the category of fault tolerance, it helps to justify the purchase of redundant components that do contribute to fault tolerance.

The following sections describe more sophisticated ways of providing server fault tolerance, beginning with server mirroring.

Server Mirroring

Server mirroring is a fault-tolerance technique in which one server duplicates the transactions and data storage of another. The servers involved must be identical machines using identical components. As you would expect, mirroring requires a link between the servers. It also entails software running on both servers that allows them to synchronize their actions continually and, in case of a failure, that permits one server to take over for the other.

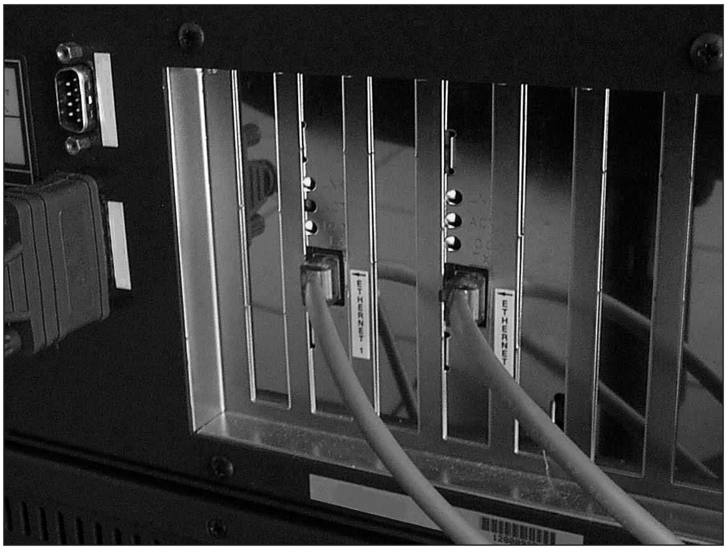


Figure 14-10 A server with redundant NICs

To illustrate the concept of mirroring, suppose that you give a presentation to a large group of people, with the audience being allowed to interrupt you to ask questions at any time.You might talk for two minutes, then wait while someone asked a question, then answer the question, then begin lecturing again, take another question, and so on. In this sense, you act like a primary server, busily transmitting and receiving information. Now imagine that your identical twin is standing in the next room and can hear you over a loudspeaker.Your twin was instructed to say exactly what you were saying as quickly as possible after you speak, but to an empty room containing only a tape recorder. Of course, your twin must listen to you before imitating you. It takes time for the twin to digest all that you're saying and repeat it, so you must slow down your lecture and your room's question-and-answer process. A mirrored server acts in much the same way. The time it takes to duplicate the incoming and outgoing data will detrimentally affect network performance if the network handles a heavy traffic load. But if you should faint during your lecture, for example, your twin can step into your room and take over for you in very short order.The mirrored server also stands ready to assume the responsibilities of its counterpart.

One advantage to mirroring is that the servers involved can stand side by side or be positioned in geographically side-by-side locations—perhaps in two different buildings of a company's headquarters, or possibly even on opposites sides of a continent. One potential disadvantage to mirroring, however, is the time it takes for a mirrored server to assume the functionality of the failed server.This delay may last 15 to 90 seconds. Obviously, this down time makes mirroring imperfect; when a server fails, users lose network service and any data in transit at the moment of the failure will be susceptible to corruption.Another disadvantage to mirroring is its toll on the network as data are copied between sites.

Examples of mirroring software include Legato System’s StandbyServer and NSI Software’s Double-Take. Although such software can be expensive, the hardware costs of mirroring are even more significant because one server is devoted to simply acting as a “tape recorder” for all data in case the other server fails. Depending on the potential cost of losing a server’s functionality for any period of time, however, the expense involved may be justifiable.

> You may be familiar with the term “mirroring” as it refers to Web sites on the Internet. Mirrored Web sites are locations on the Internet that dynamically duplicate other locations on the Internet, to ensure their continual availability. They are similar to, but not necessarily the same as, mirrored servers.

Server Clustering

Server clustering is a fault-tolerance technique that links multiple servers together to act as a single server. In this configuration, clustered servers share processing duties and appear as a single server to users. If one server in the cluster fails, the other servers in the cluster will automatically take over its data transaction and storage responsibilities. Because multiple servers can perform services independently of other servers, as well as ensure fault tolerance, clustering is more cost-effective than mirroring.

To understand the concept of clustering, imagine that you and several colleagues (who are not exactly like you) are giving separate talks in different rooms in the same conference center simultaneously. All of your colleagues are constantly aware of your lecture, and vice versa. If you should faint during your lecture, one of your colleagues can immediately jump into your spot and pick up where you left off, without the audience ever noticing. (At the same time, your colleague must continue to present his own lecture, which means that he will have to split his time between these two tasks.)

To detect failures, clustered servers regularly poll each other on the network, essentially asking,“Are you still there?”They then wait a specified period of time before again asking, “Are you still there?” If they don’t receive a response from one of their counterparts, the clustering software initiates the fail-over.This process may take anywhere from a few seconds to a minute, because all information about a failed server’s shared resources must be gathered by the cluster. Unlike with mirroring, users will not notice the switch. Later, when the other servers in the cluster detect that the missing server has been replaced, they will automatically relinquish that server’s responsibilities.The fail-over and recovery processes are transparent to network users.

One disadvantage to clustering is that the clustered servers must be geographically close—although the exact distance depends on the clustering software employed.Typically, clustering is implemented among servers located in the same data room. Some clusters can contain servers as far as a mile apart, but clustering software manufacturers recommend a closer proximity. Before implementing a server cluster, you should determine your organization’s fault-tolerance needs and fully research the options available on your servers’ platforms.

Despite its geographic limitations, clustering offers many advantages over mirroring. Each server in the cluster can perform its own data processing; at the same time, it is always ready to take over for a failed server if necessary. Not only does this ability to perform multiple functions reduce the cost of ownership for a cluster of servers, but it also improves performance.

Like mirroring, clustering is implemented through a combination of software and hardware. Novell's NetWare 5.x and Microsoft's Windows 2000 DataCenter Server and Advanced Server NOSs now incorporate options for server clustering. Clustering has been part of the UNIX operating system since the early 1990s.

Storage

Related to the availability and fault tolerance of servers is the availability and fault tolerance of data storage. In the following sections you will learn about different methods for making sure shared data and applications are never lost or irretrievable.

Redundant Array of Inexpensive Disks (RAID)

A Redundant Array of Inexpensive Disks (RAID) is a collection of disks that provide fault tolerance for shared data and applications. A group of hard disks is called a disk **array** (or a drive). The collection of disks that work together in a RAID configuration is often referred to as the "RAID drive." To the system, the multiple disks in a RAID drive appear as a single logical drive. The advantage of using RAID is that a single disk failure will not cause a catastrophic loss of data.

Although RAID comes in many different forms (or levels), all types use shared, multiple physical or logical hard disks to ensure data integrity and availability. Some RAID designs also increase storage capacity and improve performance. RAID is typically used on servers, but not on workstations because of its cost. It's important to keep in mind that RAID relies on a combination of software and hardware. The software may be a third-party package, or it may exist as part of the network operating system. On a Windows 2000 server, for example, RAID drives are configured through the Disk Management tool.

RAID Level 0 – Disk Striping. **RAID Level 0** (otherwise known as **disk striping**) is a very simple implementation of RAID in which data are written in 64 KB blocks equally across all disks in the array. Disk striping is not a fault-tolerant method because if one disk fails, the data contained in it will be inaccessible. Thus RAID Level 0 does not provide true redundancy. Nevertheless, it does use multiple disk partitions effectively, and it improves performance by utilizing multiple disk controllers. The multiple disk controllers allow several instructions to be sent to the disks simultaneously.

Figure 14-11 illustrates how data are written to multiple disks in RAID Level 0. Notice how each 64 KB piece of data is written to one discreet area of the disk array. For example, if you were saving a 128 KB file, the file would be separated into two pieces and saved in different areas of the drive. Although RAID Level 0 is easy to implement, it should not be used on mission-critical servers because of its lack of fault tolerance.

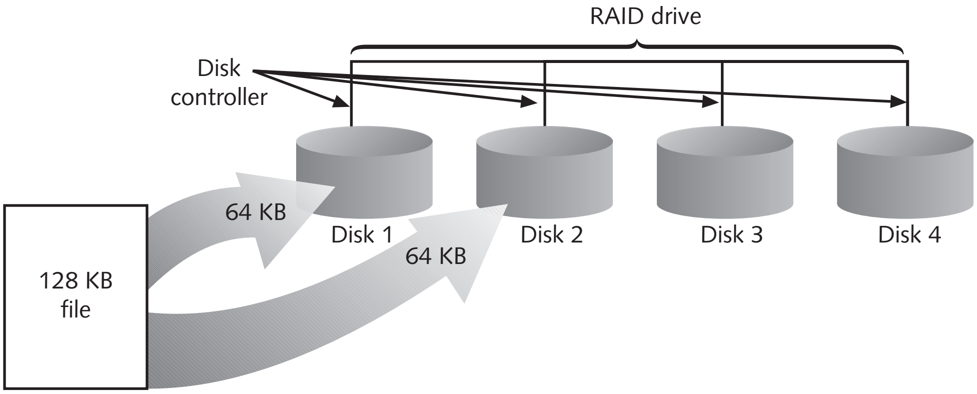


Figure 14-11 RAID Level 0 — disk striping

RAID Level 1 – Disk Mirroring. **RAID Level 1** provides redundancy through a process called **disk mirroring**, in which data from one disk are copied to another disk automatically as the information is written. Because data are continually saved to multiple locations, disk mirroring provides a dynamic data backup. If one disk in the array fails, the disk array controller will automatically switch to the disk that was mirroring the failed disk. Users will not even notice the failure. After repairing the failed disk, the network administrator must perform a resynchronization to return it to the array. As the disk's twin has been saving all of its data while it was out of service, this task is rarely difficult.

The advantages of RAID Level 1 derive from its simplicity and its automatic and complete data redundancy. On the other hand, because it requires two identical disks instead of just one, RAID Level 1 is somewhat costly. In addition, it is not the most efficient means of protecting data, as it usually relies on system software to perform the mirroring, which taxes CPU resources. Figure 14-12 depicts a 128 KB file being written to a disk array using RAID Level 1.

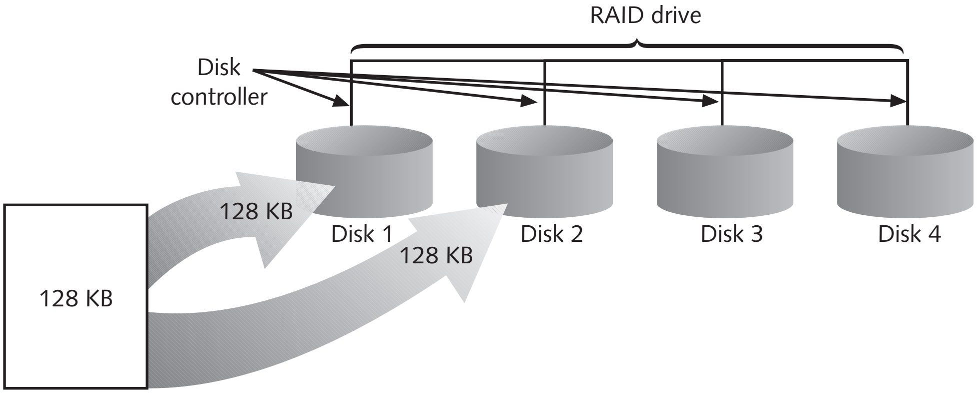


Figure 14-12 RAID Level 1 — disk mirroring

Although they are not covered in this chapter, RAID levels 2 and 4 also exist. These versions of RAID are rarely used, however, because they are less reliable or less efficient than Levels 1, 3, and 5.

RAID Level 3 – Disk Striping with Parity ECC. **RAID Level 3** involves disk striping with a special type of error correction code (ECC) known as parity error correction code. The term **parity** refers to the mechanism used to verify the integrity of data by making the number of bits in a byte sum to either an odd or even number. To accomplish parity, a parity bit (equal to either 0 or 1) is added to the bits' sum. Table 14-1 expresses how the sums of many bits achieve even parity through a parity bit. Notice that the numbers in the fourth column are all even. If the summed numbers in the fourth column were odd, an odd parity would be used. A system may use either even parity or odd parity, but not both.

Table 14-1 The use of parity bits to achieve parity

Original Data	Sum of Data Bits	Parity Bit	Sum of Data Plus Parity Bits
01110010	4	0	4
00100010	2	0	2
00111101	5	1	6
10010100	3	1	4

Parity tracks the integrity of data on a disk. It does not reflect the data type, protocol, transmission method, or file size. A parity bit is assigned to each data byte when it is transmitted or written to a disk. When data are later read from the disk, the data's bits plus the parity bit are summed again. If the parity does not match (for example, if the end sum is odd but the system uses even parity), then the system assumes that the data have suffered some type of damage. The process of comparing the parity of data read from disk with the type of parity used by the system is known as **parity error checking**.

In RAID Level 3, parity error checking takes place when data are written across the disk array. If the parity error checking indicates an error, the RAID Level 3 system can automatically correct it. The advantage of using RAID 3 is that it provides a high data transfer rate when reading from or writing to the disks. This quality makes RAID 3 particularly well suited to applications that require high speed in data transfers, such as video editing. A disadvantage of RAID 3 is that the parity information appears on a single disk, which represents a potential single point of failure in the system. Figure 14-13 illustrates how RAID Level 3 works.

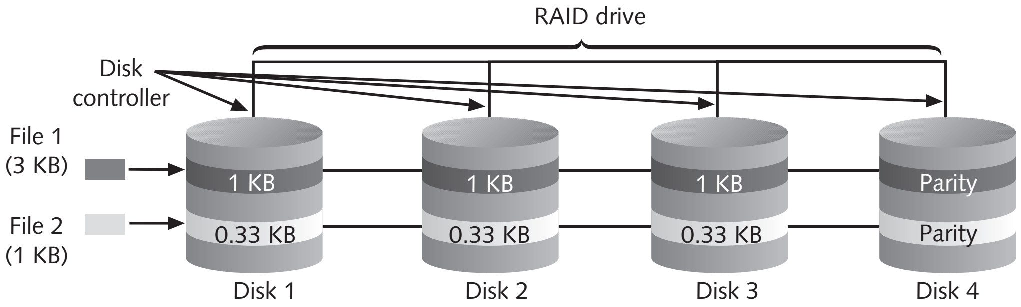


Figure 14-13 RAID Level 3 — disk striping with parity ECC

RAID Level 5 – Disk Striping with Distributed Parity. **RAID Level 5** is the most popular, highly fault-tolerant, data storage technique in use today. In RAID Level 5, data are written in small blocks across several disks. At the same time, parity error checking information is distributed among the disks, as pictured in Figure 14-14.

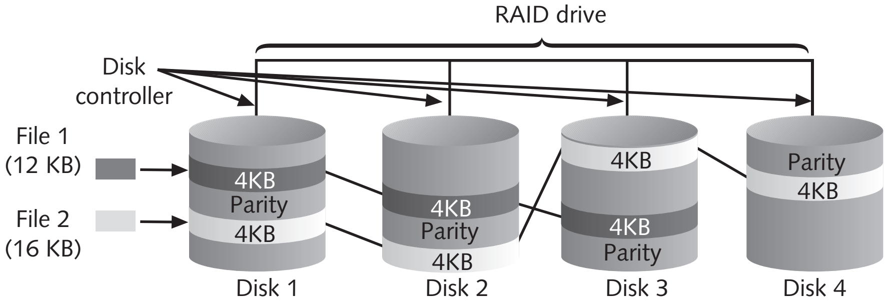


Figure 14-14 RAID Level 5 — disk striping with distributed parity

RAID Level 5 is similar to, but has several advantages over, RAID Level 3. First, it can write data more rapidly because the parity information can be written by any one of the several disk controllers in the array. Unlike RAID Level 3, RAID Level 5 uses several disks for parity information, making it more fault-tolerant. Also, RAID Level 5 allows you to replace failed disks with good ones without any interruption of service.

Network Attached Storage

Network attached storage (NAS) is a specialized storage device or group of storage devices that provides centralized fault-tolerant data storage for a network. NAS differs from RAID in that it maintains its own interface to the LAN rather than relying on a separate server to connect it to the network and control its functions. In fact, you can think of NAS as a unique type of server dedicated to data sharing. The advantage to using NAS over a typical file server is that a NAS device contains its own file system that is optimized to save and serve files (as opposed to also managing printing, authenticating login IDs, and so on). Because of this optimization, NAS reads and writes from its disk significantly faster than other types of servers could.

Another advantage to using NAS is that it can be easily expanded without interrupting service. For instance, if you purchased a NAS device with 40 GB of disk space, then six months later realized you need three times as much storage space, you could add the new 80 GB to the NAS device without requiring users to log off the network or taking down the NAS device. After physically installing the new disk space, the NAS device would recognize the added storage and add it to its pool of available reading and writing space. Compare this process to adding hard disk space to a typical server, for which you would have to take the server down, install the hardware, reformat the drive, integrate it with your NOS, then add directories, files, and permissions as necessary.

Although NAS is a separate device with its own file system, it still cannot communicate directly with clients on the network. When using NAS, the client requests a file from its usual file server (such as a Windows 2000, Linux, or NetWare 5.1 server) over the LAN. The server then requests the file from the NAS device on the network. In response, the NAS device retrieves the file and transmits it to the server, which transmits it to the client. Figure 14-15 depicts how NAS operates on a LAN.

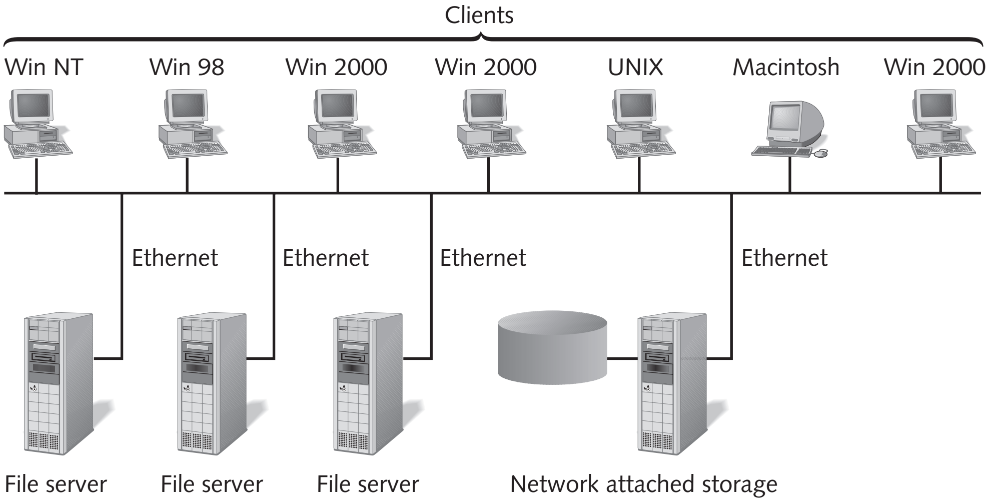


Figure 14-15 Network attached storage on a LAN

NAS is appropriate for small- or medium-sized enterprises that require not only fault tolerance, but also fast access for their data. For example, a local ISP might use NAS for hosting its customers' Web pages. Since NAS devices can store and retrieve data for any type of client (providing it can run TCP/IP), NAS is also appropriate for organizations that use a mix of different operating systems on their desktops.

The two major vendors of network attached storage are Network Appliance, Inc. and EMC Corporation. In addition, computer manufacturers such as Hewlett-Packard, Compaq and Dell now offer their own NAS solutions.

Larger enterprises that require even faster access to data and larger amounts of storage, might prefer storage area networks over NAS.You will learn about storage area networks in the following section.

Storage Area Networks

As you have learned, NAS devices are separate storage devices, but they still require a file server to interact with other devices on the network. In contrast, **storage area networks (SANs)** are distinct networks of storage devices that communicate directly with each other and with other networks. In a typical SAN, multiple storage devices are connected to multiple, identical servers. This type of architecture is similar to the mesh topology in WANs, the most fault-tolerant type of topology possible. If one storage device within a SAN suffers a fault, data is automatically retrieved from elsewhere in the SAN. If one server in a SAN suffers a fault, another server steps in to perform its functions.

Not only are SANs extremely fault tolerant, but they are also extremely fast. Much of their speed can be attributed to **Fibre Channel**, a distinct network transmission method that relies on fiber-optic media and its own, proprietary protocol. Fibre Channel connects devices within the SAN and also connects the SAN to other networks. Fibre Channel is capable of 1-Gbps (and soon, 2-Gbps) throughput. Because it depends on Fibre Channel, and not on a traditional network transmission method (for example, 10BaseT or 100BaseT), a SAN is not limited to the speed of the client/server network for which it provides data storage. In addition, since the SAN does not belong to the client/server network, it does not have to contend with the normal overhead of that network, such as broadcasts and acknowledgments. Likewise, a SAN frees the client/server network from the traffic-intensive duties of backing up and restoring data.

Figure 14-16 shows a SAN connected to a traditional Ethernet network.

Like NAS, SANs provide the benefit of being highly scalable. Once you establish a SAN, you can easily add not only further storage, but also new devices to the SAN without disrupting client/server activity on the network. Finally, SANs use a more efficient method of writing data than both NAS devices and typical client/server networks use, making them even faster.

SANs are not without drawbacks, however. One noteworthy disadvantage to implementing SANs is their high cost. A small storage area network can cost $500,000 (as much as the most expensive type of NAS) while a large SAN costs several millions of dollars. In addition, since SANs are appreciably more complex than NAS or RAID systems, investing in a SAN means also investing in long hours of training for technical staff before installation, plus significant administration efforts to keep the SAN functional.

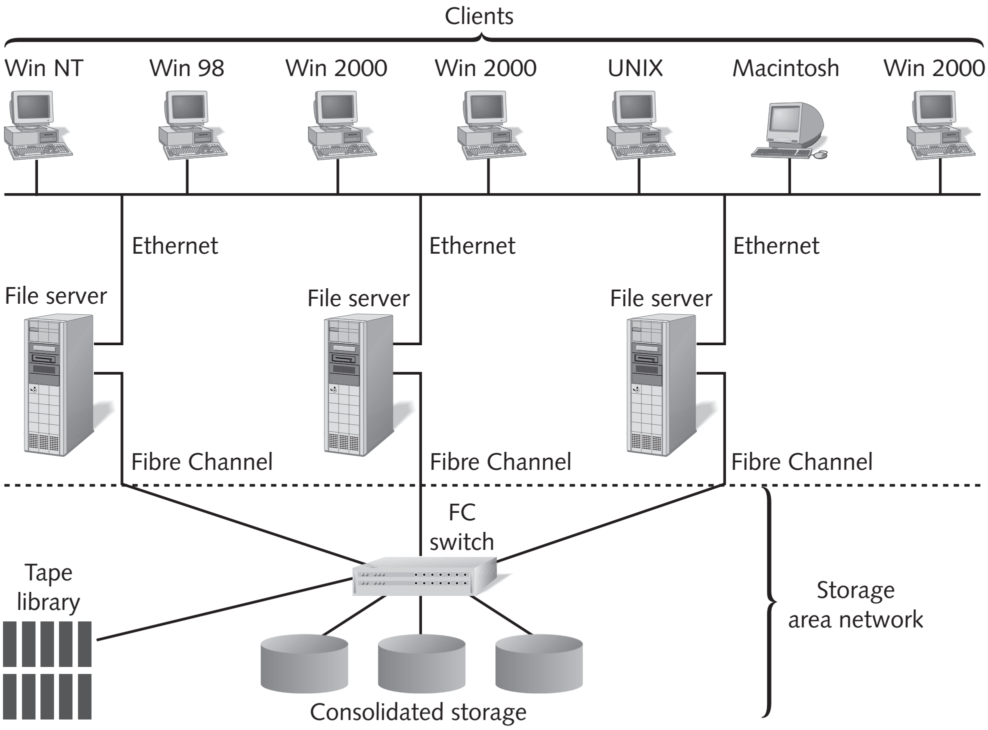


Figure 14-16 A storage area network

Because of their very high fault tolerance, massive storage capabilities and speedy data access, SANs are best suited to environments with huge quantities of data that must always be quickly available. Usually, such an environment belongs to a very large enterprise. A SAN is typically used to house multiple databases—for example, inventory, sales, safety specifications, payroll, and employee records for an international manufacturing company.

DATA BACKUP

You have probably heard or even spoken the axiom, "Make regular backups!" A **backup** is a copy of data or program files created for archiving or safekeeping purposes. Without backing up your data, you risk losing everything through a hard disk fault, fire, flood, or malicious or accidental erasure or corruption. No matter how reliable and fault-tolerant you believe your server's hard disk (or disks) to be, you still risk losing everything unless you make backups on separate media and store them off-site.

To fully appreciate the importance of backups, imagine coming to work one morning to find that everything disappeared from the server: programs, configurations, data files, user IDs, passwords, and the network operating system. It doesn't matter how it happened.

What matters at this point is how long it will take to reinstall the network operating systems; how long it will take to duplicate the previous configuration; and how long it will take to figure out which IDs should reside on the server, which groups they should belong in, and which rights each group should have. What will you say to your colleagues when they learn that all of the data that they have worked on for the last year is irretrievably lost? When you think about this scenario, you will quickly realize that you can't afford *not* to perform regular backups.

Some network administrators don't pay enough attention to backups because they find the process confusing or difficult to track. True, many different options exist for making backups. They can be performed by different types of software and hardware combinations, including via network operating system utilities. In this section, you will learn about the most common methods of performing data backup, ways to schedule them, and methods for determining what you need to back up. Backup methods unsuitable for large systems, such as floppy disks or other removable storage media, are not covered in this section. Note that backing up workstations and backing up servers and other host systems are different operations. To qualify for Net+ certification, you should focus on making server backups.

Tape Backups

Currently, the most popular method for backing up networked systems is tape backup, because this method is simple and relatively economical. Tape backups require the use of a tape drive connected to the network (via a system such as a file server or dedicated, networked workstation), software to manage and perform backups, and, of course, backup media. The tapes used for tape backups resemble small cassette tapes, but they are of a higher quality, specially made to reliably store data. Figure 14-17 depicts two types of backup tape media: 4 mm and 8 mm.

On a relatively small network, standalone tape drives may be attached to each server. On a large network, one large, centralized tape backup device may manage all of the subsystems' backups. This tape backup device will usually be connected to a computer other than a busy file server to reduce the possibility that backups might cause traffic bottlenecks. Extremely large environments (for example, global manufacturers with several terabytes of inventory and product information to safeguard) may require robots to retrieve and circulate tapes from a tape storage library (or **vault**) that may be as large as a warehouse. Figure 14-18 illustrates how tape drives typically fit into a medium or large network.

IMATION DDS-90
4mm DATA TAPE/BANDE DE DONNÉES
Media Recognition System
Digital Data Storage
IMATION
DDS-90
4mm DATA TAPE/BANDE DE DONNÉES
IMATION 3M
8mm DATA TAPE / BANDE DE DONNÉES
D8-112
112M

Figure 14-17 Examples of backup tape media

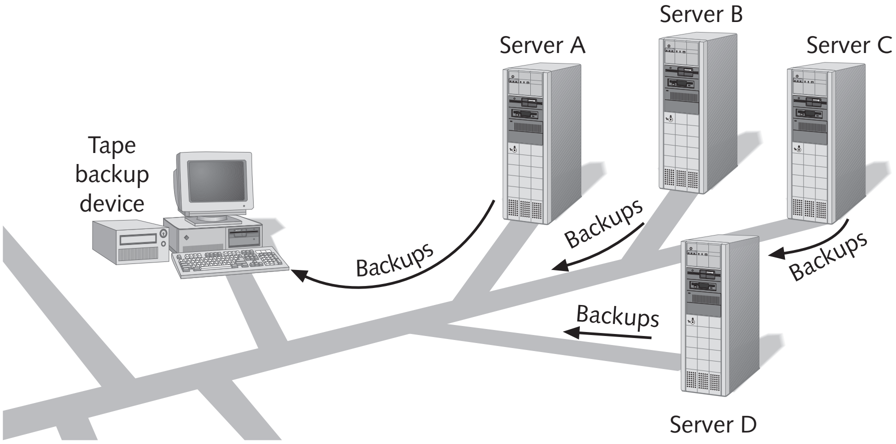


Figure 14-18 A tape drive on a medium or large network

To select the appropriate tape backup solution for your network, you should consider the following questions:

- Does the backup drive or media provide sufficient storage capacity?
- Are the backup software and hardware proven to be reliable?

- Does the backup software use data error checking techniques?
- Is the system quick enough to complete the backup process before daily operations resume?
- How much do the tape drive, software, and media cost?
- Will the backup hardware and software be compatible with existing network hardware and software?
- Does the backup system require frequent manual intervention? (For example, will staff members need to become involved in tape rotation?)
- Will the backup hardware, software, and media accommodate your network's growth?

Examples of tape backup software include Computer Associates' ARCserve, Dantz Development Corporation's Retrospect, Hewlett-Packard's Colorado and OmniBack, IBM's ADSTAR Distributed Storage Manager (ADSM), NovaStor Corporation's NovaNET, and Veritas Software Corporation's Backup Exec. Popular tape drive manufacturers include Exabyte, Hewlett-Packard, IBM, Quantum, Seagate, and Sony.You will need to consult the software and hardware specifications to determine whether a particular backup system is compatible with your network.

Online Backups

Many companies on the Internet now offer to back up data over the Internet—that is, to perform **online backups**. Usually, online backup providers require you to install their client software.You also need a connection to the Internet. Online backups implement strict security measures to protect the data in transit, as the information must traverse public carrier links. Most online backup providers allow you to retrieve your data at any time of day or night, without calling a technical support number. Both the backup and restoration processes are entirely automated. In case of a disaster, the online backup company may offer to create CD-ROMs containing your servers' data.

A potential drawback to online backups is that the cost of this service can vary widely. In addition, despite strict security controls, it may be difficult to verify that your data has been backed up successfully. Online backup providers include @Backup, Atrieva, Connected, HotWired, and Safeguard.

When evaluating an online backup provider, you should test its speed, accuracy, security, and, of course, the ease with which you can recover the backed up data. Be certain to test the service before you commit to a long-term contract for online backups.

Backup Strategy

After selecting the appropriate tool for performing your servers' data backups, you should devise a backup strategy to guide you and your colleagues in performing reliable backups that provide maximum data protection. This strategy should be documented in a common area (for example, on a Web site accessible to all IT staff) and should address at least the following questions:

- What kind of rotation schedule will backups follow?
- At what time of day or night will the backups occur?
- How will you verify the accuracy of the backups?
- Where will backup media be stored?
- Who will take responsibility for ensuring that backups occurred?
- How long will you save backups?
- Where will backup and recovery documentation be stored?

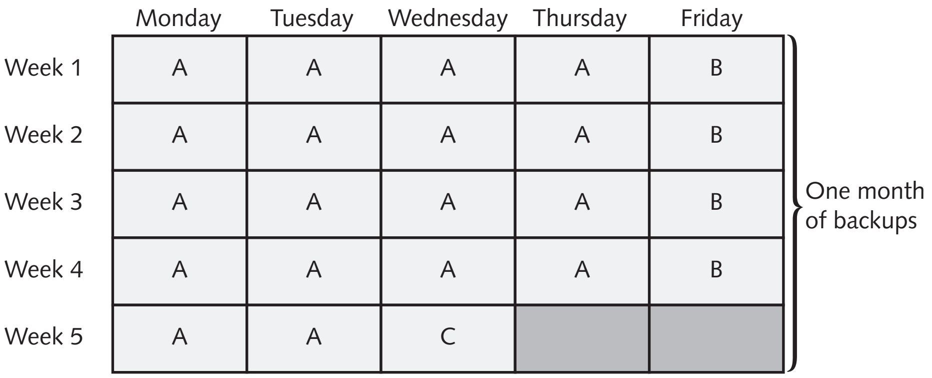
Different backup methods provide varying levels of certainty and corresponding labor and cost. The various methods are described below:

- **Full backup**—All data on all servers are copied to a storage medium, regardless of whether the data are new or changed.
- **Incremental backup**—Only data that have changed since the last backup are copied to a storage medium.
- **Differential backup**—Only data that have changed since the last backup are copied to a storage medium, and that information is then marked for subsequent backup, regardless of whether it has changed.

When managing network backups, you need to determine the best possible **backup rotation scheme**—that is, you need to create a plan that specifies when and how often backups will occur. The aim of a good backup rotation scheme is to provide excellent data reliability without overtaxing your network or requiring a lot of intervention. For example, you might think that backing up your entire network's data every night is the best policy because it ensures that everything is completely safe. But what if your network contains 50 GB of data and is growing by 10 GB per month? Would the backups even finish by morning? How many tapes would you have to purchase? Also, why should you bother backing up files that haven't changed in three weeks? How much time will you and your staff need to devote to managing the tapes? How would the transfer of all of the data affect your network's performance? All of these considerations point to a better alternative than the "tape-a-day" solution—that is, an option that promises to maximize data protection but reduce the time and cost associated with backups.

When planning your backup strategy, you can choose from several standard backup rotation schemes. The most popular of these schemes, called **grandfather-father-son**, uses daily (son), weekly (father), and monthly (grandfather) backup sets. As depicted in Figure 14-19, in the grandfather-father-son scheme, three types of backups are performed each month: daily incremental (every Monday through Thursday), weekly full (every Friday), and monthly full backups (last day of the month).

In this scheme, backup tapes are reused regularly. For example, week 1's Monday tape would also serve as week 2's and week 3's Monday tape. One day each week, a full backup, called "father," is recorded in place of an incremental one and labeled for the week to which it corresponds—for example, "week 1," "week 2," and so on. This "father" tape is reused monthly—for example, October's week 1 tape would be reused for November's week 1 tape. The final set of media is labeled "month 1," "month 2," and so on, according to which month of the quarter the tapes will be used. This "grandfather" medium records full backups on the last business day of each month and is reused quarterly. Each of these media may consist of a single tape or a set of tapes, depending on the amount of data involved. A total of 12 media sets are required for this basic rotation scheme, allowing for a history of two to three months.



	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	A	A	A	A	B
Week 2	A	A	A	A	B
Week 3	A	A	A	A	B
Week 4	A	A	A	A	B
Week 5	A	A	C	▨	▨

A = Incremental "son" backup (daily)
B = Full "father" backup (weekly)
C = Full "grandfather" backup (monthly)

Figure 14-19 The grandfather-father-son backup rotation scheme

Once you have determined your backup rotation scheme, you should ensure that backup activity is recorded in a backup log. Information that belongs in a backup log include the backup date, tape identification (day of week or type), type of data backed up (for example, Accounting Department spreadsheets or a day's worth of catalog orders), type of the backup (full, incremental, or differential), files that were backed up, and site at which the tape is stored. Having this information available in case of a server failure will greatly simplify data recovery.

Finally, once you begin to back up network data, you should establish a regular schedule of verification. In other words, from time to time (depending on how often your data change and how critical the information is), you should attempt to recover some critical files from your backup media. Many network administrators can attest that the darkest hour of their career was when they were asked to retrieve critical files from a backup tape and found that no backup data existed because their backup system never worked in the first place!

DISASTER RECOVERY

Disaster recovery is the process of restoring your critical functionality and data after an enterprise-wide outage that affects more than a single system or a limited group of users. Disaster recovery must take into account the possible extremes, rather than relatively minor outages, failures, security breaches, or data corruption. In a disaster recovery plan, you should consider the worst-case scenarios, from a far-reaching hurricane to a military attack. You should also consider what might happen if your typical networking staff isn't available. The plan should outline multiple contingencies, in case your best options don't pan out. Although you must attend to all of the protection methods discussed in this chapter, disaster recovery also requires a comprehensive strategy for restoring functionality and data after things go terribly awry.

Every organization should have a disaster recovery team (with an appointed coordinator) and a disaster recovery plan. This plan should address not only computer systems, but also power, telephony, and paper-based files. When writing the sections of the plan related to computer systems, your team should specifically address the following issues:

- Contact names for emergency coordinators who will execute the disaster recovery response in case of disaster, as well as roles and responsibilities of other staff.
- Details on which data and servers are being backed up, how frequently backups occur, where backups are kept (off-site), and, most importantly, how backed up data can be recovered in full.
- Details on network topology, redundancy, and agreements with national service carriers, in case local or regional vendors fall prey to the same disaster.
- Regular strategies for testing the disaster recovery plan.
- A plan for managing the crisis, including regular communications with employees and customers. Consider the possibility that regular communications modes (such as phone lines) might be unavailable.

Having a comprehensive disaster recovery plan not only lessens the risk of losing critical data in case of extreme situations, but also makes potential customers and your insurance providers look more favorably on your organization.

CHAPTER SUMMARY

- Integrity refers to the soundness of your network's files, systems, and connections. To ensure their integrity, you must protect them from anything that might render them unusable, such as corruption, tampering, natural disasters, and viruses. Availability of a file or system refers to how consistently and reliably it can be accessed by authorized personnel.

- Several basic measures can be employed to protect data and systems on a network: (1) prevent anyone other than a network administrator from opening or changing the system files; (2) monitor the network for unauthorized access or changes; (3) record authorized system changes in a change management system; (4) install redundant components; (5) perform regular health checks on the network; (6) monitor system performance, error logs, and the system log book regularly; (7) keep backups, boot disks, and emergency repair disks current and available; and (8) implement and enforce security and disaster recovery policies.

- A virus is a program that replicates itself so as to infect more computers, either through network connections or through floppy disks passed among users. Viruses may damage files or systems or simply annoy users by flashing messages or pictures on the screen or by causing the computer to beep.

- Many other unwanted and potentially destructive programs are mistakenly called viruses. For example, a program that disguises itself as something useful but actually harms your system is called a Trojan horse. An example of a Trojan horse is an executable file sent to you over the Internet that purportedly installs a new game, but actually reformats your hard disk.

- Boot sector viruses are the most common types of viruses. They reside on the boot sector of a floppy disk and become transferred to the partition sector or the DOS boot sector on a hard disk. The only way a boot sector virus can move from a floppy to a hard disk is if the floppy disk is left in the drive when the machine starts up.

- Macro viruses take the form of a word-processing or spreadsheet program macro, which may be executed when you use the word-processing or spreadsheet program. Macro viruses were the first type of virus to infect data files rather than executable files. Because data files are more apt to be shared among users and because macro viruses are typically easier to write than executable viruses, these viruses have quickly become widespread.

- File-infected viruses attach themselves to executable files. When the infected executable file runs, the virus copies itself to memory. Later, the virus will attach itself to other executable files.

- Network viruses take advantage of network protocols, commands, messaging programs, and data links to propagate themselves. Although all viruses could theoretically travel across network connections, network viruses are specially designed to take advantage of network vulnerabilities.

- Worms are not technically viruses, but rather programs that run independently and travel between computers and across networks. Although they do not alter other programs as viruses do, worms may carry viruses.

- Any type of virus may have additional characteristics that make it harder to detect and eliminate. These characteristics may be encrypted, stealth, polymorphic, or time-dependent.

- Although a well-written virus attempts to avoid detection, you may suspect the presence of a virus on your system if you notice any of the following symptoms: unexplained increases in file sizes; programs (such as Microsoft Word) launching, running, or exiting more slowly than usual; unusual error messages appearing without probable cause; significant, unexpected loss of system memory; or fluctuations in display quality.

- A good antivirus program should be able to detect viruses through signature scanning, integrity checking, and heuristic checking. It should also be compatible with your network environment, centrally manageable, easy to use (transparent to users), and not prone to false alarms.

- Antivirus software is merely one piece of the puzzle in protecting your network from viruses. An antivirus policy is another essential component. It should provide rules for using antivirus software and policies for installing programs, sharing files, and using floppy disks. Furthermore, it should be authorized and supported by the organization's management and should include sanctions for disobeying the policy.

- A virus hoax is a false alert about a dangerous, new virus that could seriously damage your workstation. Virus hoaxes usually have no realistic basis and should be ignored.

- In broad terms, a failure is a deviation from a specified level of system performance for a given period of time. A fault, on the other hand, is the malfunction of one component of a system. A fault can result in a failure. The goal of fault-tolerant systems is to prevent faults from progressing to failures.

- Fault tolerance is a system's capacity to continue performing despite an unexpected hardware or software malfunction. It can be achieved in varying degrees, with the optimal level of fault tolerance for a system depending on how critical its services and files are to productivity. At the highest level of fault tolerance, a system will be unaffected by a drastic problem, such as a power failure.

- An excellent way to achieve fault tolerance is to provide duplicate elements to compensate for faults in critical components, a practice known as redundancy. You can implement redundancy for servers, cabling, routers, hubs, gateways, NICs, hard disks, power supplies, and other components.

- To assess the fault tolerance of your network you must look for single points of failure—places on the network where, if a fault occurs, the transfer of data may break down without possibility of an automatic recovery.

- As you consider sophisticated fault-tolerance techniques for servers, routers, and WAN links, remember to address the environment in which your devices operate. Protecting your data also involves protecting your network from excessive heat or moisture, break-ins, and natural disasters.

- Networks cannot tolerate power loss or less than optimal power.You will have to guard against the following power flaws: blackouts, brownouts (sags), surges, and line noise.
- A UPS is a battery-operated power source directly attached to one or more devices and to a power supply (such as a wall outlet), which prevents undesired features of the power source from harming the device or interrupting its services. UPSs vary widely in the type of power aberrations they can rectify, the length of time they can provide power, and the number of devices they can support.
- A standby UPS provides continuous voltage to a device by switching virtually instantaneously to the battery when it detects a loss of power from the wall outlet. Upon restoration of the power, the standby UPS switches the device to use A/C power again. A standby UPS requires a brief service outage when it detects that A/C power has stopped; in this time, a sensitive device (such as a server) may have already detected the power loss and shut down or restarted.
- An online UPS uses the A/C power from the wall outlet to continuously charge its battery, while providing power to a network device through its battery. In other words, a server connected to an online UPS always relies on the UPS battery for its electricity. An online UPS provides the best kind of power redundancy available. Because the server never needs to switch from the wall outlet's power to the UPS's power, no risk of momentarily losing service exists.
- To choose the best UPS for your network, you must consider a number of factors: the amount of power needed, the period of time in which you must keep a device running, line conditioning, and cost.
- If your organization cannot withstand a power loss, either because of its computer services or other electrical needs, you might consider investing in an electrical generator for your building. Generators can be powered by diesel, liquid propane gas, natural gas, or steam. They do not provide surge protection, but they do provide clean (free from noise) electricity.
- The type of network topology that offers the best fault tolerance is a mesh topology. In a mesh network, nodes are connected either directly or indirectly by multiple pathways. In a mesh topology, data can travel over these multiple paths from any one point to another.
- The physical media you use may also offer redundancy. A SONET ring, for example, can easily recover from a fault in one of its links because it forms a self-healing ring.
- When components are hot swappable, they have identical functions and can automatically assume the functions of their counterpart if it suffers a fault. They are called hot swappable because they can be changed (or swapped) while a machine is still running (hot).
- The use of multiple components enables load balancing, or an automatic distribution of traffic or processing to optimize response.
- As with other devices, you can make servers more fault-tolerant by supplying them with redundant components. Critical servers often contain redundant NICs, processors, and/or hard disks. These redundant components provide assurance that if one fails, the whole system won't fail, and they enable load balancing.

- A fault-tolerance technique that involves utilizing a second, identical server to duplicate the transactions and data storage of one server is called server mirroring. Mirroring can take place between servers that are either geographically side by side or distant. Mirroring requires not only a link between the servers, but also software running on both servers to enable the servers to continually synchronize their actions and to permit one to take over in case the other fails.
- Server clustering is a fault-tolerance technique that links multiple servers together to act as a single server. In this configuration, clustered servers share processing duties and appear as a single server to users. If one server in the cluster fails, the other servers in the cluster will automatically take over its data transaction and storage responsibilities.
- An important server redundancy feature is a Redundant Array of Inexpensive Disks (RAID). All types of RAID use shared, multiple physical or logical hard disks to ensure data integrity and availability; some designs also increase storage capacity and improve performance. RAID is typically used on servers, but not on workstations because of its added cost. RAID is accomplished through a combination of both software and hardware.
- RAID Level 0 is a very simple implementation of RAID in which data are written in 64 KB blocks equally across all of the disks in the array, a technique known as disk striping. Disk striping is not a fault-tolerant method because if one disk fails, the data contained in it will be inaccessible. Thus RAID Level 0 does not provide true redundancy.
- RAID Level 1 provides redundancy through a process called disk mirroring, in which data from one disk are automatically copied to another disk as the information is written. This option can be considered a dynamic data backup. If one disk in the array fails, the disk array controller will automatically switch to the disk that was mirroring the failed disk.
- RAID Level 3 involves disk striping with parity error correction code. Parity refers to the integrity of the data as expressed in the number of 1s contained in each group of correctly transmitted bits. In RAID Level 3, parity error checking takes place when the data are written across the disk array.
- RAID Level 5 is the most popular, highly fault-tolerant, data storage technique in use today. In RAID Level 5, data are written in small blocks across several disks; parity error checking information is also distributed among the disks.
- Network attached storage (NAS) is a device or group of devices attached to a client/server network dedicated to data storage. It uses its own file system but relies on a traditional network transmission method such as Ethernet to interact with the rest of the client/server network.
- A storage area network (SAN) is a distinct network of multiple storage devices and servers that provides fast, highly available, and highly fault-tolerant access to large quantities of data for a client/server network. SAN uses a proprietary network transmission method (such as Fibre Channel) rather than a traditional network transmission method such as Ethernet.

- A backup is a copy of data or program files created for archiving or safekeeping purposes. If you do not back up your data, you risk losing everything through a hard disk fault, fire, flood, or malicious or accidental erasure or corruption. No matter how reliable and fault-tolerant you believe your server's hard disk (or disks) to be, you still risk losing everything unless you make backups on separate media and store them off-site.
- Currently, the most popular method for backing up networked systems is tape backup, because it is simple and relatively economical. Tape backups require a tape drive connected to the network (via a system such as a file server or dedicated, networked workstation), software to manage and perform backups, and backup media.
- To select the appropriate tape backup solution for your network, you should consider the following issues: storage capacity; proven reliability; data error checking techniques; speed; cost of the tape drive, software, and media; compatibility with existing network hardware and software; and extent of automation.
- Many companies on the Internet now offer to back up data over the Internet—that is, to perform online backups. Usually, online backup providers require that you have their client software in addition to a connection to the Internet. They implement strict security measures to protect the data in transit, because the information must traverse public carrier links. Both the backup and restore processes are entirely automated.
- A good backup strategy should be well documented and should address at least the following questions: What kind of rotation schedule will backups follow? At what time of day or night will the backups occur? How will you verify the accuracy of backups? Where will backup media be stored? Who will take responsibility for ensuring that backups occurred? How long will you save backups? Where will backup and recovery documentation be stored?
- Different backup methods provide varying levels of certainty and corresponding labor and cost. A full backup copies all data on all servers to a storage medium, regardless of whether the data are new or changed. An incremental backup copies only data that have changed since the last backup A differential backup copies only data that have changed since the last backup, and that information is marked for subsequent backup, regardless of whether it has changed.
- If you are responsible for the network's backups, your most important decision will relate to the backup rotation scheme. The aim of a good backup rotation scheme is to provide excellent data reliability but not to overtax your network or require much intervention.
- The most popular backup rotation scheme is called "grandfather-father-son." This scheme uses daily (son), weekly (father), and monthly (grandfather) backup sets.
- Once you have determined your backup rotation scheme, you should ensure that backup activity is recorded in a backup log. Information that belongs in a backup log include the following: when the backup took place; which tape was used (day of week or type); which data were backed up; whether the backup was full, incremental, or differential; which files were backed up; and where the tape is stored. Having this information available in case of a server failure will greatly simplify data recovery.

- Disaster recovery is the process of restoring your critical functionality and data after an enterprise-wide outage that affects more than a single system or a limited group of users. It must account for the possible extremes, rather than relatively minor outages, failures, security breaches, or data corruption. In a disaster recovery plan, you should consider the worst-case scenarios, from a hurricane to a military attack.
- Every organization should have a disaster recovery team (with an appointed coordinator) and a disaster recovery plan. The plan should address not only computer systems, but also power, telephony, and paper-based files.

KEY TERMS

array — A group of hard disks.

availability — How consistently and reliably a file, device, or connection can be accessed by authorized personnel.

backup — A copy of data or program files created for archiving or safekeeping purposes.

backup rotation scheme — A plan for when and how often backups occur, and which backups are full, incremental, or differential.

blackout — A complete power loss.

boot sector virus — A virus that resides on the boot sector of a floppy disk and is transferred to the partition sector or the DOS boot sector on a hard disk. A boot sector virus can move from a floppy to a hard disk only if the floppy disk is left in the drive when the machine starts up.

brownout — A momentary decrease in voltage, also known as a *sag*. An overtaxed electrical system may cause brownouts, recognizable as a dimming of the lights.

differential backup — A backup method in which only data that have changed since the last backup are copied to a storage medium, and that information is marked for subsequent backup, regardless of whether it has changed.

disaster recovery — The process of restoring critical functionality and data to a network after an enterprise-wide outage that affects more than a single system or a limited group of users.

disk mirroring — A RAID technique in which data from one disk are automatically copied to another disk as the information is written.

disk striping — A simple implementation of RAID in which data are written in 64 KB blocks equally across all disks in the array.

encrypted virus — A virus that is encrypted to prevent detection.

fail-over — The capability for one component (such as a NIC or server) to assume another component's responsibilities without manual intervention.

failure — A deviation from a specified level of system performance for a given period of time. A failure occurs when something doesn't work as promised or as planned.

fault — The malfunction of one component of a system. A fault can result in a failure.

fault tolerance — The capacity for a system to continue performing despite an unexpected hardware or software malfunction.

Fibre Channel — A distinct network transmission method that relies on fiber-optic media and its own, proprietary protocol. Fibre Channel is capable of 1-Gbps (and soon, 2-Gbps) throughput.

file-infected virus — A virus that attaches itself to executable files. When the infected executable file runs, the virus copies itself to memory. Later, the virus will attach itself to other executable files.

full backup — A backup in which all data on all servers are copied to a storage medium, regardless of whether the data are new or changed.

grandfather-father-son — A backup rotation scheme that uses daily (son), weekly (father), and monthly (grandfather) backup sets.

hard disk redundancy — See *Redundant Array of Inexpensive Disks (RAID).*

heuristic scanning — A type of virus scanning that attempts to identify viruses by discovering "virus-like" behavior.

hot swappable — A characteristic that enables identical components to be interchanged (or swapped) while a machine is still running (hot). Once installed, hot swappable components automatically assume the functions of their counterpart if it suffers a fault.

incremental backup — A backup in which only data that have changed since the last backup are copied to a storage medium.

integrity — The soundness of a network's files, systems, and connections. To ensure integrity, you must protect your network from anything that might render it unusable, such as corruption, tampering, natural disasters, and viruses.

integrity checking — A method of comparing the current characteristics of files and disks against an archived version of these characteristics to discover any changes. The most common example of integrity checking involves a checksum.

intrusion detection — The process of monitoring the network for unauthorized access to its devices.

line noise — Fluctuations in voltage levels caused by other devices on the network or by electromagnetic interference.

load balancing — An automatic distribution of traffic over multiple links, hard disks, or processors intended to optimize responses.

macro viruses — A newer type of virus that takes the form of a word-processing or spreadsheet program macro, which may execute when a word-processing or spreadsheet program is in use.

network attached storage (NAS) — A device or set of devices attached to a client/server network that is dedicated to providing highly fault-tolerant access to large quantities of data. NAS depends on traditional network transmission methods such as Ethernet.

network virus — A type of virus that takes advantage of network protocols, commands, messaging programs, and data links to propagate itself. Although all viruses could theoretically travel across network connections, network viruses are specially designed to attack network vulnerabilities.

online backup — A technique in which data are backed up to a central location over the Internet.

online UPS — A power supply that uses the A/C power from the wall outlet to continuously charge its battery, while providing power to a network device through its battery.

parity — The mechanism used to verify the integrity of data by making the number of bits in a byte sum to either an odd or even number.

parity error checking — The process of comparing the parity of data read from a disk with the type of parity used by the system.

polymorphic virus — A type of virus that changes its characteristics (such as the arrangement of its bytes, size, and internal instructions) every time it is transferred to a new system, making it harder to identify.

RAID Level 0 — An implementation of RAID in which data are written in 64 KB blocks equally across all disks in the array.

RAID Level 1 — An implementation of RAID that provides redundancy through disk mirroring, in which data from one disk are automatically copied to another disk as the information is written.

RAID Level 3 — An implementation of RAID that uses disk striping for data and parity error correction code on a separate parity disk.

RAID Level 5 — The most popular, highly fault-tolerant, data storage technique in use today, RAID Level 5 writes data in small blocks across several disks. At the same time, it writes parity error checking information among several disks.

redundancy — The use of more than one identical component for storing, processing, or transporting data.

Redundant Array of Inexpensive Disks (RAID) — A server redundancy measure that uses shared, multiple physical or logical hard disks to ensure data integrity and availability. Some RAID designs also increase storage capacity and improve performance. See also *disk striping,* and *disk mirroring*.

sag — See *brownout*.

server clustering — A fault-tolerance technique that links multiple servers together to act as a single server. In this configuration, clustered servers share processing duties and appear as a single server to users. If one server in the cluster fails, the other servers in the cluster will automatically take over its data transaction and storage responsibilities.

server mirroring — A fault-tolerance technique in which one server duplicates the transactions and data storage of another, identical server. Server mirroring requires a link between the servers and software running on both servers so that the servers can continually synchronize their actions and take over in case the other fails.

signature scanning — The comparison of a file's content with known virus signatures (unique identifying characteristics in the code) in a signature database to determine whether the file is a virus.

standby UPS — A power supply that provides continuous voltage to a device by switching virtually instantaneously to the battery when it detects a loss of power from the wall outlet. Upon restoration of the power, the standby UPS switches the device to use A/C power again.

stealth virus — A type of virus that hides itself to prevent detection. Typically, stealth viruses disguise themselves as legitimate programs or replace part of a legitimate program's code with their destructive code.

storage area network (SAN) — A distinct network of multiple storage devices and servers that provides fast, highly available, and highly fault-tolerant access to large quantities of data for a client/server network. SAN uses a proprietary network transmission method (such as Fibre Channel) rather than a traditional network transmission method such as Ethernet.

surge — A momentary increase in voltage due to distant lightning strikes or electrical problems.

time-dependent virus — A virus programmed to activate on a particular date. This type of virus, also known as a "time bomb," can remain dormant and harmless until its activation date arrives.

Trojan horse — A program that disguises itself as something useful but actually harms your system.

uninterruptible power supply (UPS) — A battery-operated power source directly attached to one or more devices and to a power supply (such as a wall outlet), which prevents undesired features of the power source from harming the device or interrupting its services.

vault — A large tape storage library.

virus — A program that replicates itself so as to infect more computers, either through network connections or through floppy disks passed among users. Viruses may damage files or systems or simply annoy users by flashing messages or pictures on the screen or by causing the keyboard to beep.

virus hoax — A rumor, or false alert, about a dangerous, new virus that could supposedly cause serious damage to your workstation.

volt-amp (VA) — A measure of electrical power. A volt-amp is the product of the voltage and current (measured in amps) of the electricity on a line.

worm — An unwanted program that travels between computers and across networks. Although worms do not alter other programs as viruses do, they may carry viruses.

REVIEW QUESTIONS

1. Describe five scenarios that might detrimentally affect the integrity or availability of your network's data.
2. Which of the following percentages represents the highest availability for a network?
    a. 0.10%
    b. 0.01%
    c. 99%
    d. 99.99%

3. To ensure that a system change does not detrimentally affect integrity and availability, what information should you record about the change?
    a. who performed the change and why it was necessary
    b. when the change occurred, why it was necessary, who performed the change, and what the change involved
    c. what the change involved and when it occurred
    d. when the change occurred and how to reverse it
4. Which of the following symptoms might make you suspect that your workstation is infected with a macro virus?
    a. Your computer takes a long time to start up.
    b. While in Microsoft Word, you receive a message that says, "WXYC rules the roost."
    c. While navigating through folders, your icons suddenly switch from pictures of folders to pictures of pineapples.
    d. You can no longer save word-processing files to your hard disk.
5. Why are stealth viruses difficult to detect?
    a. They attach themselves to legitimate programs.
    b. They frequently change their file size characteristics.
    c. They disguise themselves as legitimate programs.
    d. They destroy the file allocation table to prevent directory scanning.
6. Name three key components of an enterprise-wide antivirus policy.
7. Which of the following is a popular antivirus program?
    a. Norton VirusPro
    b. Scandisk
    c. Norton AntiVirus
    d. McAfee Virex
8. A worm is a type of polymorphic virus. True or False?
9. How does a Trojan horse disguise itself?
    a. It frequently changes its code characteristics.
    b. It disguises itself as a useful program.
    c. It prevents the user from performing directory scans.
    d. It does not appear in a directory listing.
10. Which of the following techniques does a polymorphic virus employ to make itself more difficult to detect?
    a. It frequently changes its code characteristics.
    b. It disguises itself as a useful program.

   c. It damages the file allocation table to prevent directory scanning.
   d. It moves from one location to another on the hard disk.

11. If your antivirus software uses signature scanning, what must you do to keep its virus-fighting capabilities current?
    a. Purchase new virus signature scanning software every three months.
    b. Reinstall the virus scanning software each month.
    c. Manually edit the signature scanning file.
    d. Regularly update the antivirus software's signature database.
12. What might you tell a user who receives what seems to be a virus hoax message?
    a. Ignore and delete the message.
    b. Open the message to verify that it is indeed a hoax.
    c. Send the message to the help desk.
    d. Save the message for you to review.
13. Describe the main difference between a fault and a failure.
14. Fail-over is a technique used in highly fault-tolerant systems. True or False?
15. What makes two components hot swappable?
    a. Both are similar and installed in the same device.
    b. Both are similar and one can be quickly swapped in for the other in case of a fault.
    c. Both are identical, both are installed in the same device, and one can instantly take over from the other in case of a fault.
    d. Both are identical and one can be quickly swapped in for the other in case of a fault.
16. Over time, what might electrical line noise do to your system?
    a. wear down the power switch
    b. damage the internal circuit boards
    c. increase the system board's response time
    d. cause more frequent outages
17. How long will an online UPS take to switch its attached devices to battery power?
    a. 15 seconds
    b. 10 seconds
    c. 5 seconds
    d. no time

18. Which of the following is the most highly fault-tolerant network topology?
    a. bus
    b. ring
    c. partial mesh
    d. full mesh
19. Which characteristic of SONET rings makes them highly fault-tolerant?
    a. They are self-healing.
    b. They are geographically diverse.
    c. They are made of fiber-optic cable.
    d. They share traffic over many lines.
20. Describe how load balancing between redundant NICs works.
21. Why is simple disk striping not fault-tolerant?
    a. It can be performed only on a single disk drive.
    b. If one disk fails, data contained on that disk are unavailable.
    c. It does not keep a dynamic record of where data are striped.
    d. It relies on a single disk controller.
22. Why is RAID Level 5 superior to RAID Level 3?
23. Which of the following can be considered an advantage of server clustering over server mirroring?
    a. Clustering does not affect network performance.
    b. Clustering fail-over takes place more rapidly.
    c. Clustering has no geographical distance limitations.
    d. Clustering keeps a more complete copy of a disk's data.
24. What is currently the greatest disadvantage to using server clustering?
    a. It's expensive.
    b. It detrimentally affects performance.
    c. It requires that servers in a cluster be geographically close.
    d. It is difficult to maintain.
25. List four considerations that you should weigh when deciding on a data backup solution.

26. Which factor must you consider when using online backups that you don't typically have to consider when backing up to a LAN tape drive?
    a. reliability
    b. geographical distance
    c. security
    d. time to recover
27. In a grandfather-father-son backup scheme, the October–week 1–Thursday backup tape would contain what types of files?
    a. files changed since last Thursday
    b. files changed since a month ago Thursday
    c. files changed since Wednesday
    d. files changed since a week ago Wednesday
28. Which of the following is a major disadvantage to performing full system backups on a daily basis?
    a. They would take too long to perform.
    b. They would take too long to restore.
    c. They would be less reliable than incremental backups.
    d. They would require manual intervention.
29. How can you verify the accuracy of tape backups?
30. Name four components of a smart disaster recovery plan.

HANDS-ON PROJECTS

In the following Hands-on Projects, you will have a chance to experiment with some fault-tolerance measures. Bear in mind that solutions will vary with each network environment.

Project 14-1

For this project, you will need a NetWare 5.x server with a Windows 2000 Professional client workstation attached. The server should contain at least a Pentium processor, 70 MB of RAM, 100 MB of free disk space, in addition to the NetWare operating system (with all the latest patches) and its connection to the network. The client workstation should contain at least a Pentium processor, 64 MB of RAM, 200 MB of free disk space, a CD-ROM drive, and the Novell Client for NetWare. You should be able to connect to not only the NetWare server, but also the Internet from that workstation. You will also need a copy of the Norton AntiVirus Corporate Edition for NetWare Servers software on CDs.

To install Norton AntiVirus on a NetWare server:

1. Log onto the server as an administrator from the Windows 2000 Professional workstation attached to your NetWare server.

2. Map a drive to the server's **SYS** volume.
3. Insert the Norton AntiVirus CD number 2 into your workstation's CD-ROM drive. If the CD menu does not automatically open, open **My Computer**, double-click the CD-ROM drive, then double-click on the **setup.exe** file to begin the installation process.
4. Select the **Install Norton AntiVirus to Servers** option, then click **Next**. The License Agreement dialog box appears.
5. After reading the license agreement, check **I agree**, then click **Next** to continue. The Select Items dialog box appears.
6. Check the **Server Program** option, then click **Next** to continue.
7. Double-click **NetWare Services**. The Select Computers dialog box appears.
8. Double-click **NetWare Directory Services**, then select the SYS volume object where you want to install the AntiVirus software. (To navigate through the NDS tree, double-click the tree object, then select organizational units until you find the one that contains the SYS volume object you want.) Click **Add**.
9. Enter the appropriate container name, and the user name and password for this container, as prompted, then click **Next** to continue.
10. You will be prompted for a location to install the Norton AntiVirus program. Keep the default install path and click **Next** to continue. The Select Server Group dialog box appears.
11. Type the name **CLASS** for the new server group and click **Next** to continue. You will be asked to confirm that you want to create this server group. Click **Yes** to confirm.
12. Select **Manual Startup** and click **Next** to continue. The Using The Symantec System Center Program dialog box appears.
13. Click **Next** until you reach the final Setup screen, reading each screen of instructions carefully, then click **Close**. The AntiVirus installation commences.
14. Now that you have installed the software on the server, you will need to initialize it. At the server console, type **load sys:nav\vpstart.nlm /install** to initialize the Norton AntiVirus program. After the NLM has loaded, you can use Norton AntiVirus on your NetWare server to detect viruses.
15. At the Windows 2000 Professional workstation, experiment with the NAV program to immediately scan servers, change configuration options, and set a regularly scheduled server scan.

Project 14-2

Because the Norton AntiVirus software uses signature scanning as one of its antivirus measures, you will have to update the signature database on a regular schedule. In this exercise, you will update the software you installed on your NetWare server in Project 14-1.

1. Open your Web browser and go to **www.sarc.com/avcenter/download.html**, the Symantec Security Updates page.

2. Click **Download Virus Definitions Updates** in the center of the page. The Download Virus Definitions page opens.
3. Use the list arrow to select **English, US** (if it is not already selected).
4. From the list of Symantec products, click **Norton AntiVirus for NetWare**.
5. Click **Download Updates** to continue. The Download English Updates page appears.
6. Click on the name of the update file that is appropriate for Norton AntiVirus Corporate Edition (the version you installed in Project 14-1). The File Download dialog box opens.
7. Click **OK** to choose to save the file to disk.
8. The Save As dialog box opens. Save the file to your **C:\TEMP** (or a similar temporary) directory. Click **Save** to begin the download.
9. If you aren't already logged in as Administrator, log onto the network from your workstation as an Administrator. Run the file you just downloaded, supplying the location of your NAV NLM.
10. Follow instructions on the screen to ensure that your antivirus signature database was updated.

Project 14-3

In this exercise, you will use an online UPS capacity tool to determine the UPS needed for an imaginary network server. UPS vendors such as APC supply these online tools so that you do not have to calculate by hand the VA necessary for your network. To complete this project, you will need a workstation with access to the Internet.

1. From the networked workstation, launch the Web browser and go to **www.apcc.com/template/size/apc/**. This Size UPS Web site provides a UPS sizing utility that you can use to determine your UPS capacity needs. In this case, you want to determine the needs of your server.
2. Click the **Server** link in the middle of the screen. The UPS Selector page opens. In the middle of the screen, a drop-down list of server types appears.
3. Click the list arrow, click **Compaq ProLiant 850R**, then click **Submit**. A configuration page for this server opens, allowing you to specify a number of options that characterize your server.
4. To the default specifications, add a 22-inch LCD monitor, one attached tape drive, and four external hard drives.
5. Click **Add to Configuration**. A new page opens, allowing you to set more parameters for your server.
6. Click **Continue to User Preferences**.
7. Note the defaults, including a 20-minute run time.

8. Choose your region from the drop-down list to make sure the correct voltage is used in the UPS power requirements calculation.
9. Click **Show Solution**. A new page opens, containing the recommendations for the type of server that you specified.
10. Scroll to the bottom of the page to view your configuration. How many volts did the utility estimate your configuration would require? How many watts and VA would the UPS have to supply to keep the server, monitor, tape drive, and external hard disks running for 20 minutes?
11. Click the **Back** button on your browser to return to the last set of parameters you specified.
12. Click the **Delete Device** button near the bottom of the page to erase the configuration you have just generated.
13. Click **OK** to confirm that you want to delete this device from your list of UPS configurations.
14. The UPS Selector Web page appears. Click **Add Another Device**.
15. Repeat Steps 2 through 8, this time selecting an EMC Celera SE server. How many volts would this configuration require? How many watts and VA would the UPS have to supply to keep the Celera SE server running for 20 minutes?

CASE PROJECTS

Case Project

1. You have been asked to help a local hospital improve its network's fault tolerance. The hospital's network carries critical patient care data in real time from both a mainframe host and several servers to PCs in operating rooms, doctors' offices, the billing office, teaching labs, and remote clinics located across the region. Of course, all of the data transferred is highly confidential and must not be lost or accessed by unauthorized personnel. Specifically, the network consists of the following:

   - Six hundred PCs are connected to five shared servers that run Novell NetWare 5.0. Fifty of these PCs serve as training PCs in medical school classrooms. Two hundred PCs sit in doctors' offices and are used to view and update patient records, submit accounting information, and so on. Twenty PCs are used in operating rooms to perform imaging and for accessing data in real time. The remaining PCs are used by administrative staff.

   - The PCs are connected in a mostly switched, star-wired bus network using Ethernet 100BaseTX technology. Where switches are not used, some hubs serve smaller workgroups of administrative and physician staff.

   - An Internet gateway supports e-mail, online medical searches, and VPN communications with four remote clinics. The Internet connection is a single T1 link to a local Internet service provider.

- ❒ A firewall prevents unauthorized access from the T1 connection into the hospital's network.

The hospital's IT director has asked you to identify the critical points of failure in her network and to suggest how she might eliminate them. On a sheet of paper, draw a logical diagram of the network and identify the single points of failure, then recommend which points of failure should be addressed to increase availability and how to achieve this goal. For each fault-tolerant component or method you recommend, find manufacturers' data available on the Web to identify its cost.

2. Unfortunately, the solution you provided for the hospital was rejected by the board of directors because it was too expensive. How would you determine where to cut costs in the proposal? What questions should you ask the IT director? What points of failure do you suggest absolutely must be addressed with redundancy?
3. Your second proposal, with its reduced cost, was accepted by the board of directors. Now the hospital's IT director has asked you to outline a disaster recovery plan. Based on what you have learned about the hospital's topology, usage patterns, and current fault-tolerance measures, develop a disaster recovery plan for the hospital that specifically addresses how functionality and data will be restored.
4. After you submitted your outline of the hospital's disaster recovery plan, the IT director takes you aside and confesses that she isn't sure whether her network administrator is doing the right thing with the hospital's antivirus software and policy. Currently, the antivirus software is installed on each workstation in the hospital and scans each workstation's memory and hard disk once per week. She asks whether you have a solution for a better antivirus implementation and whether she should ask users to scan their hard disks more frequently than once per week. How do you respond?